Centre interuniversitaire de recherche sur les réseaux d'entreprise, la logistique et le transport

Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation

Multi-Modal Scheduled Service Network Design with Resource Management for Two-Tier City Logistics

Pirmin Fontaine
Teodor Gabriel Crainic
Ola Jabali
Walter Rei

Juin 2017

CIRRELT-2017-27

Bureaux de Montréal :
Université de Montréal
Pavillon André-Aisenstadt
C.P. 6128, succursale Centre-ville
Montréal (Québec)
Canada H3C 3J7
Téléphone : 514 343-7575
Télécopie : 514 343-7121

Bureaux de Québec :
Université Laval
Pavillon Palasis-Prince
2325, de la Terrasse, bureau 2642
Québec (Québec)
Canada G1V 0A6
Téléphone : 418 656-2073
Télécopie : 418 656-2624

www.cirrelt.ca

Multi-Modal Scheduled Service Network Design with Resource Management for Two-Tier City Logistics

Pirmin Fontaine[1], Teodor Gabriel Crainic[1,*], Ola Jabali[3], Walter Rei[1]

[1] Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT) and Department of Management and Technology, Université du Québec à Montréal, P.O. Box 8888, Station Centre-Ville, Montréal, Canada H3C 3P8

[2] Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Piazza Leonardo da Vinci, 32, Milano 20133, Italy

Abstract. We introduce a new problem for selecting services in a tactical plan of a two-tier city logistics system. Compared to existing models, we consider different transportation modes with multiple compartments within one model. Moreover, we integrate inbound and outbound demands in the model and consider the used resources in the constraints. For defining the problem, we introduce a new integer programming formulation that has many similarities to the well-known knapsack and bin packing problem. To efficiently solve this problem, we use this formulation in a Benders decomposition algorithm which is implemented in a Branch-and-Cut framework. We further improve the method by using pareto-optimal cuts, partial decomposition and valid inequalities. The numerical results show significant run time improvements compared to Cplex and illustrate the benefits of including both flows and different transportation modes with multiple compartments in the planning.

Keywords: City logistics, Benders decomposition, network design, tactical planning.

Acknowledgements. We gratefully acknowledge the financial support provided by the Natural Sciences and Engineering Council of Canada (NSERC), through its Discovery and Acceleration Grant programs. We also gratefully acknowledge the support of Fonds de recherche du Quebec through their Strategic Clustering Grant program, the Canadian Foundation for Innovation and the Ministère de l'Économie, de la Science et de l'Innovation of Quebec through the Leaders Fond Infrastructure Grant program, as well as Calcul Québec and Compute Canada for access to their high-performance computing infrastructure. While working on this project, the second author was also Adjunct Professor, Department of Computer Science and Operations Research, Université de Montréal.

---

* Corresponding author: teodorgabriel.crainic@cirrelt.net

Dépôt légal – Bibliothèque et Archives nationales du Québec
Bibliothèque et Archives Canada, 2017

1 Introduction

The transportation of goods in urban areas is a complex activity. This complexity is amplified by the increasing urbanization. In 2014, 54% of the world's population was living in urban areas. The United Nations (2014) are expecting an increase up to 66% until 2050 and the OECD 85% until 2100. This results not only in larger demands but also attributes to increasing the intricacy of distributions networks. Increasing volume of e-commerce further strains distribution systems. In Europe, B2C commerce increased in the last years by more than 13% each year and in 2016 an increase of 12% is expected as well. To improve the distribution, new organizations and business models are developed under the topic of city logistics (CL). Hereby, city logistics aims to reduce congestion and pollution in the city, which are both even more amplified through urbanization and increasing e-commerce. Integrating goods distribution with existing public transport infrastructure is one way to improve quality of living in cities and reduce congestion. By consolidating flows in and also out of the city, CL can improve the utilization of the means of transport and reduce driving distances. This also leads to a more sustainable transportation network.

Due to its increasing importance, CL found more attention during the last years. Bektas et al. (2015) and Savelsbergh and Van Woensel (2016) recently reviewed the CL literature and pointed out the potential for future research. The current literature mostly considers a two echelon system where the goods are consolidated outside of the city. Then they are transported to smaller redistribution points in the city where the goods are loaded into smaller vehicles for the final distribution. However, these models only consider one vehicle type for the first echelon which is assumed to be a truck. More green transportation modes like subway or tram systems are not yet considered, even though practical test with cargo trams in several European cities showed the potential of these modes. Further, current models only consider flows into the city. This results, however, in empty truck movements out of the city that could be used for outbound demand. To the best of our knowledge, also resources have not been addressed explicitly in planning models for city logistics. Finally, the literature for solving these planning models efficiently is scare.

We want to fill these gaps in the literature, by introducing a tactical planning model for the operated services in the first echelon of system. Our model considers both inbound and as well outbound demands. We further consider in our problem setting not only trucks but also scheduled services, which can be for example trams. While the strength of the truck is the flexibility in routing, a tram takes benefits from its multiple compartments which are essential for the combination of inbound and outbound demand. We evaluate how a CL system can benefit when combining different transportation modes, especially under the aspect of inbound and outbound traffic. Moreover, we model the resources, like fleet size and satellite capacity, in our model.

We introduce a new integer programming formulation for the problem. This formulations benefits from the fact that, compared to classical network design formulations, we do not need to consider the paths in the network. To solve larger instances, we develop an efficient Benders decomposition algorithm. The numerical results show significant computational benefits of our approach compared to solving the integer programming formulation with Cplex. Furthermore, we show that considering different transportation modes and combining inbound and outbound flows in one model is important and should be considered.

The contributions of this paper are as follows: (1) We define a two-tier city logistics model for in- and outbound logistics which considers different transportation modes with multiple compartments and resource management, (2) we propose an efficient Benders decomposition algorithm based on a new formulation for solving the problem to optimality, (3) we demonstrate the efficiency of the proposed solution method through extensive numerical experiments, and (4) we show the benefits of combining inbound and outbound demand in one model and of using multi-model fleet structures.

The remainder of the paper is structured as follows: In Section 2, we define the problem and relate it to the existing literature. The mathematical model is introduced in Section 3. Then, Section 4 introduces the Benders decomposition algorithm and the different improvement strategies which are used. Numerical results on the computational performance and managerial insights are given in Section 5. The paper concludes with a summary and an outlook for future research directions.

2 Problem Setting and Related Literature

We consider a two-tier city logistics (2T-CL) system as proposed by Crainic et al. (2004). In Section 2.1, we recall the introduced system and basic vocabulary. The considered problem is defined in Section 2.2. Including resource constraints into the model is one of the main contributions of this paper. Therefore, we will describe the different resources, we are looking at, in Section 2.3.

2.1 2T-CL Sytem

As shown in Figure 1, we consider a two-tier city logistics (2T-CL) system as proposed by Crainic et al. (2009). The 2T-CL system consists of two layers with the goal to deliver goods from origin to destination (in the classical setting from outside the city to customers in the city).

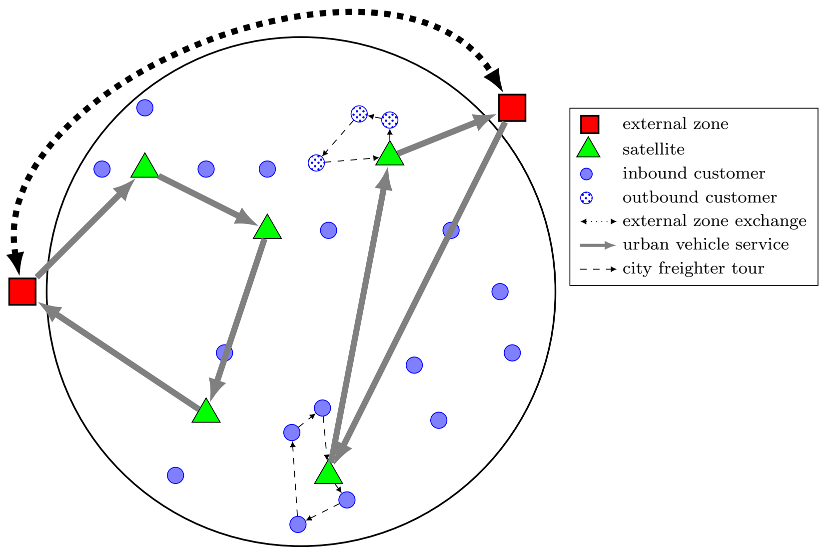

Figure 1: Two-tier City Logistics System

All arriving goods have to be shipped to consolidation centers (logistics platforms, urban distribution centers) at the border of the city. We will refer to them as *external zones*. Hereby, goods can be transferred between external zones by using for example a ring tram. After consolidation and sorting in the final external zones, the goods are shipped via *urban vehicles* to so called *satellites* from where the final distribution to the *customers* is done by *city freighters*.

The demands are spread over the city and either have to be delivered from outside the city to the customer location or are picked up at the customer location and transported to an external zone. Customers can further specify a time-window for the delivery/pick-up at its location and also a time-window at the external zone for the long-haul connection. While all proposed models only assume inbound traffic, Crainic et al. (2012) point out that the integration of outbound traffic and traffic within the city is important but also arises further algorithmic challenges which are however not solved in the paper by Crainic et al. (2012).

2.2 Problem Definition and Related literature

Most publications investigate two-echelon Vehicle Routing Problems (e.g., Perboli et al., 2011; Hemmelmayr et al., 2012). Instead, Crainic et al. (2009) introduce a general framework for two-tier city logistics (2T-CL) systems and use a service network design problem for the first tier. Since we also focus on a service network design formulation on the first tier, the interested reader is referred to Bektas et al. (2015) for a detailed literature analysis on Vehicle Routing in CL.

Although service network design is a rather new topic in the area of CL, it is applied in freight transportation (Crainic, 2000) or other settings like bike sharing (Neumann-Saavedra et al., 2016) for a while. Recently, also the consideration of resources found more interest: Andersen et al. (2009b,a) are the first publications considering resources (or here called assets) in Service Network Design. Based on these publications, Crainic et al. (2014) enlarge the range of this aspect within tactical planning models. These authors further include location decisions (Crainic et al., 2015b) and resource acquisition for scheduled services (Crainic et al., 2016b). Since these models are already very difficult to solve in deterministic case, Wang et al. (2016) analyzed the value of the deterministic solution for stochastic network design problems.

The urban vehicles are typically modeled as trucks with larger capacity which still have a flexibility in timing and routing. We explicitly want to consider and distinguish between scheduled services and free transportation modes. While the later one are trucks with a defined capacity, scheduled service can be trams, subway systems or trains which have fixed schedules and special compartments for the transportation of goods (Lindholm and Behrends, 2012). Paris, for example, is considering a system, which we define as scheduled services: In the proposed system, loading and unloading of trams has to be performed on secondary stations. Therefore, a special freight tram has to be scheduled within the passenger transport schedule. This gives also more flexibility since these stations can be used for interim parking. The integration of passenger and freight transportation, was so far only investigated for scheduled passenger service (Ghilas et al., 2013), but not in the context of city-logistics.

The more complex urban vehicle structure, also leads to a larger variety of satellites. Not only parking lots or free spaces in the city for trucks can be used; also subway, rail or tram stations. However, satellites with passenger traffic will have strong limitations on space. Moreover, time limitations for unloading and loading operations arise since passenger traffic shall not be interfered.

Besides the inbound traffic, we are also considering outbound traffic. Goods can be delivered from customers (which are now the origin) via satellites to external zones (the destination). Since goods can only be loaded and unloaded to vehicles according to the LIFO rule, the ordering of goods becomes relevant. But especially for urban vehicles a higher utilization is achieved by reducing empty trips outside of the city.

Goods might have different restrictions for the shipment: For example, hazardous materials should not be shipped together with food. Or some goods even have special requirements like refrigerated compartments. Using the idea of physical internet (Crainic and Montreuil, 2016), we reduce the problem to a one product case. The different requirements are all achieved by smart boxes which can be combined for the shipping. Each good has an origin, a destination, a volume and time windows at potential external zones, as well as at the customer.

At the satellites, the goods are transferred into city freighters, which can be small trucks, cars, electric vehicles or bikes, to perform the final delivery to the customers. In the literature, this problem is modeled as a synchronized, scheduled, multidepot, multiple-tour, heterogeneous vehicle routing problem with time windows (SS-MDMT-VRPTW).

The goal of our problem is, however, to plan a schedule of first-tier services which is then repeatedly operated over the considered time horizon (for example 6 month). As suggested by Crainic et al. (2009), one possibility to decompose the problem, we approximate the routing costs of the second layer. Because of demand uncertainty and a very high detail level in the second layer, customers will have a predefined subset of possible satellites, but the actual routing does not have to be considered in the tactical planning process. Therefore, we define a service network design problem for the first layer. The time windows at the customer are transferred to the satellites by an approximation of the delivery time of the SS-MDMT-VRPTW. The actual second-tier routing of demands has to be solved on the day before when the true demands are realized. For doing so, Crainic and Sgalambro (2014) give insights in the day-before planning and decompose the problem into two easier problems. Crainic et al. (2015a) include demand uncertainty into the two-tier model and introduce strategies for adjusting the plan after demand is realized.

2.3 Resources

In CL, many capacity restrictions are important to consider. Crainic et al. (2009) define already several restrictions like vehicle or satellite capacity which we are now extending.

To ensure a smooth transfer of goods at satellites, urban vehicles and freighters are not allowed to wait at satellites in general. Each satellite is defined by a limited *space*, which can be used for transferring goods. This space restricts the size of tram or the number of urban vehicles at a parking lot at a certain time. However, no space is available for storing goods at satellites. Moreover, satellites can also have a *time* limit during which the unloading or loading has to be performed. This is especially the case if passenger services are combined with transportation of goods. A passenger tram should not wait longer at a station because of unloading or loading activities.

We assume a heterogeneous fleet for the urban vehicles: not only different means of transport also different sizes of, for example, trucks or trams. Each of the urban vehicle types has a *fleet size* at each external zone which is available over the whole planning horizon. They have a predefined *vehicle capacity* for shipping goods and a home depot while repositioning could be further considered.

Also the vehicle might be available, especially in scheduled or fixed services, the *network availability* is important. It, however, can also be interesting to restrict truck traffic in specific hours of the day.

Some of the resources (fleet size, vehicle capacity) are available over the whole planning horizon, while the resources space, time, network availability, and operators can vary over day and time. For example, during rush hour a subway system has typically no free capacity for additional freight transportation.

3 Mathematical Model

In this section, we first introduce the used notation. In Section 3.2, we formulate the problem as integer program which is ultimately used in Section 4 for the Benders decomposition.

3.1 Notation

In this model, we are optimizing a cyclic schedule for the first-tier services and the corresponding resources. The resulting plan is built for the whole schedule length, which is, depending on the problem, half a day or even shorter. This plan is then used for several month. Only the assignment of goods to the services and the second tier have to be optimized when the actual demand is realized.

The considered schedule length is divided into $t = 1, \dots, T$ periods, where each period represents a small time interval. As in (Crainic et al., 2009), the period length is defined such that (1) at most one departure of a service from its external zone may take place during a period, and, (2) all considered time-related parameters are integer multiples of the period length.

The set of external zones $\mathcal{E}$ are the facilities where the goods are loaded into the urban vehicles and transported to the set of satellites $\mathcal{Z}$. However, external zones are not considered to be satellites in our model, thus customers close to the external zones are not considered in our model. These customers are assumed to be identified in a pre-processing, but can still directly be delivered by city freighters from external zones. The external zones are divided into external zones for the different means of transport $\mathcal{M}$ (trucks, trams and subways) $\mathcal{E}^m$ (analogously for satellites $\mathcal{Z}^m$). If different modes are available in an external zone $e \in \mathcal{E}$, e is part of all relevant subsets. The same hold for satellites.

Let $\mathcal{T}_m$ be the set of available urban-vehicle types of mode m. Then let $\mathcal{T}$ be the set of all urban-vehicle types. The corresponding capacity of each urban-vehicle type is given by u_τ and the fleet size at each external zone e by $n_{e\tau}$. For trucks we assume that there is only one compartment, for trams and subways the number of same sized compartments is equal to the number of doors. The compartment capacity is given by u^c_τ and the number of compartments is n^c_τ (such that $u^c_\tau n^c_\tau = u_\tau \forall \tau \in \mathcal{T}$). Each urban-vehicle has a unique depot $e_\tau \in \mathcal{E}$, where it starts and finishes its operation at beginning and end of the planning horizon. We note that a vehicle type can be assigned to several external zones, in which case the vehicle type is duplicated.

We consider three different types of capacities at the satellites: Each satellite has a capacity $u^{\mathcal{T}}_{zt}$ denoting the number of urban vehicles for each period t which satellite z may accommodate. Moreover, for each satellite z the capacity of urban vehicles of transportation mode $m \in \mathcal{M}$ is limited to u^m_{zt} in period t. For trucks this is the actual number of trucks, while for trams or subways the number of cars can also be the limitation. This allows the model to use, for example, tram stops also as truck satellites, as in the concept of Paris. The number of goods which can be unloaded or uploaded is typically reflected by the number of city freighters which can stop at a satellite. We assume that the total volume of goods assigned to satellite z in period t cannot exceed u^V_{zt}.

The goal of the model is to satisfy the demand, given by set $\mathcal{D}$. Further, to distinguish inbound and outbound logistics, the set of demands is divided into the two disjunctive sets $\mathcal{D}^I$ for e2c and $\mathcal{D}^O$ for c2e demand. Each demand is specified by volume v_d and the customer location. A fixed cost of f_{de} is applied for assigning demand $d \in \mathcal{D}$ to external zone $e \in \mathcal{E}$. These costs avoid the exclusive assignment of products to external zones, and can be seen as costs for using a ring tram in the surrounding of the city or further transportation costs for the carrier. The costs for the best located external zone from the carrier perspective might be even zero. In the first-tier, the destination of inbound demand $d \in \mathcal{D}^I$ (or the origin of outbound demand $d \in \mathcal{D}^O$) is not the customer but a satellite. To avoid exclusive assignments as well, a set of potential satellites $\mathcal{Z}(d) \subset \mathcal{Z}$ is given with corresponding final distribution (or pickup) costs s_{dzt} for demand d from satellite z in period t. Besides operating and handling costs also costs for disturbance through freight activities during the operation time of the service are included. The same strategy with corresponding cost structures is applied for outbound traffic as well.

Moreover, each demand can have a time-window: when it will be available at the origin $[a^o_d, b^o_d]$, and when it can be delivered to the destination $[a^d_d, b^d_d]$. If products are for example available at a certain time in an external zone, but do not have a deadline until they have to leave the depot, these bounds are set to T (or in the other case to 0).

The travel-time between two points i, j in the network is defined by g_{ij}. The service time for loading and unloading an urban-vehicle τ is h_τ.

The goal is to select a set of urban vehicle *services* $\mathcal{R}$. Service r starts at external zone $e_r \in \mathcal{E}$, visits several satellites, and returns to the same external zone. The ordered set of visited satellites is given by $\sigma^r = \{z_i^r \in \mathcal{Z}, i = 1, \ldots, |\sigma^r|$, such that if r visits satellite i before satellite j then $i < j\}$. For a given service r, an urban vehicle of type τ_r, which is of mode m_r, operates service r and has the associated costs are given by k_r. The costs include again not only the operating costs of the route and unloading or loading activities, but also disturbance factors for freight activities during the operation time of the service.

The departure time of service r at its origin e_r is t_0^r. Then the urban vehicle arrives in period $t_1^r = t_0^r + g_{e_r,z_1^r}$ at the first satellite $z_1^r \in \sigma^r$. Since we also allow the waiting at specific satellites (for example of trams), the waiting time of service r at satellite z is denoted by w_{zr}. After performing loading and/or unloading operations and possible waiting, the urban vehicle leaves the first satellite at $t_1^r + h_{\tau_r} + w_{z_1^r,r}$. Thus the general schedule of service r is given by the ordered set $\{t_i^r, i = 0, 1, \ldots, |\sigma^r| + 1\}$, where $t_i^r = t_{i-1}^r + g_{z_{i-1}^r,z_i^r} + h_{\tau_r} + w_{z_i^r,r}$, representing the period the service visits satellite $z_i^r \in \sigma^r$, and the service finishes its route at the external e_r in period $t_{|\sigma^r|+1}^r$. Moreover, the fact, that a tram track is used by public transportation or a satellite is not available in specific time slots, is already included in set $\mathcal{R}$.

To reflect the different compartments of the services r, $\mathcal{R}^C(r)$ defines the set of services, where each element reflects a service of a compartment. For services with one compartment, obviously $|\mathcal{R}^C(r)| = 1$. If the service $r \in \mathcal{R}$ is operated, also all compartment services are operated. The modeling of services per compartment is necessary to prohibit the simultaneous assignment of inbound and outbound flows to the same compartment of a service at the same time.

A summary of the notation is also given in Table 7 in Appendix 1.

3.2 Mathematical formulation

Crainic et al. (2009) formulated the problem as path-based formulation by defining all possible itineraries for each demand explicitly. As also stated by the authors, the path-based formulation seems to be beneficial in a column-generation based approach. We, instead, introduce a new formulation which has many similarities with the knapsack and bin packing problems. Therefore, good bounds and efficient solution methods can be used within the solution framework.

For this formulation, we use the same binary decision variable ρ_r that indicates if the urban vehicle service $r \in \mathcal{R}$ is selected or not. However, the set of itineraries is not necessary. Instead, we introduce the decision variable $x_{r^c,d,z} = 1$, if demand $d \in \mathcal{D}$ is assigned to compartment service $r_c \in \mathcal{R}^C$ and satellite $z \in \mathcal{Z}$.

Figure 2 illustrates the new decision variable $x_{r^c,d,z}$.

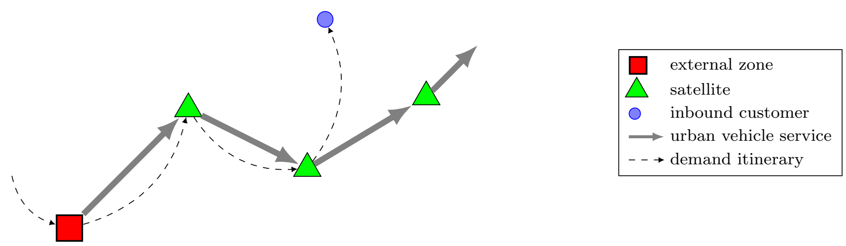

Figure 2: Example: itinerary vs. assignment decision

The demand itinerary is arriving at a time t_1 in the external zone. There it is assigned to the service (we assume here only one compartment) and shipped over satellite 1 to satellite 2, where it arrives at time $t_2 > t_1$. From the it's transported via city freighter to the customer location. Since each service is assigned to exactly one external zone, the r_c also dictates the external zone for the itinerary of demand d. Due to the fixed schedule of each service and by assigning d to one satellite z, also the time of the service at satellite z and the external zone are given. The information of the flow of the itinerary is not necessary.

Using the introduced notation and the defined decision variables, we can formulate the problem as follows:

$$\min \sum_{r \in \mathcal{R}} k_r \rho_r + \sum_{d \in \mathcal{D}} \sum_{r \in \mathcal{R}} \sum_{r^c \in \mathcal{R}^C(r)} \sum_{z \in \mathcal{Z}} \left(s_{d,z,t} + f_{d,e_r} \right) x_{r^c,d,z} \tag{1}$$

subject to

$$\sum_{z\in\mathcal{Z}}\sum_{r\in\mathcal{R}}\sum_{r^c\in\mathcal{R}^C(r)} x_{r^c,d,z} = 1 \qquad \forall d \in \mathcal{D} \tag{2}$$

$$x_{r^c,d_1,z_1} + x_{r^c,d_2,z_2} \leq 1 \qquad \forall r^c \in \mathcal{R}^C(r), r \in \mathcal{R}, d_1 \in \mathcal{D}^I, d_2 \in \mathcal{D}^O, z_1, z_2 \in \mathcal{Z}(r), z_1 \geq z_2 \tag{3}$$

$$\sum_{z\in\mathcal{Z}}\sum_{d\in\mathcal{D}^I} v_d x_{r^c,d,z} \leq u^c_{\tau_r}\rho_r \qquad \forall r^c \in \mathcal{R}^C(r), r \in \mathcal{R} \tag{4}$$

$$\sum_{z\in\mathcal{Z}}\sum_{d\in\mathcal{D}^O} v_d x_{r^c,d,z} \leq u^c_{\tau_r}\rho_r \qquad \forall r^c \in \mathcal{R}^C(r), r \in \mathcal{R} \tag{5}$$

$$\sum_{r\in\mathcal{R}(t,\tau,e)} \rho_r \leq n_{e\tau} \qquad \forall \tau \in \mathcal{T}, e \in \mathcal{E}, t = 1,\ldots,T \tag{6}$$

$$\sum_{t^-=t-h_\tau+1}^{t}\sum_{r\in\mathcal{R}(z,t^-)} \rho_r \leq u^\tau_{zt} \qquad \forall z \in \mathcal{Z}, t = 1,\ldots,T \tag{7}$$

$$\sum_{t^-=t-h_\tau+1}^{t}\sum_{r\in\mathcal{R}(z,t^-,m)} \rho_r \leq u^m_{zt} \qquad \forall z \in \mathcal{Z}, m \in \mathcal{M}, t = 1,\ldots,T \tag{8}$$

$$\sum_{r\in\mathcal{R}(t,z)}\sum_{r^c\in\mathcal{R}^C(r)}\sum_{d\in\mathcal{D}} v_d x_{r^c,d,z} \leq u^V_{zt} \qquad \forall z \in \mathcal{Z}, t = 1,\ldots,T \tag{9}$$

$$\rho_r \in \{0,1\} \qquad \forall r \in \mathcal{R} \tag{10}$$

$$x_{r^c,d,z} \in \{0,1\} \qquad \forall d \in \mathcal{D}, r^c \in \mathcal{R}^C(r), r \in \mathcal{R}, z \in \mathcal{Z} \tag{11}$$

The objective function (1) is minimizing the costs of selecting and operating a service, plus the costs for assigning a demand to a satellite, plus the costs for assign a demand to an external zone. The assignment costs include the operational costs at that terminal as well as the transportation costs from and to the satellite/external zone.

Constraint (2) ensures that each item is assigned exactly to one compartment. Equation (3) ensures that outbound demand is only assigned to a compartment after the inbound demand is loaded off and the compartment is empty. The capacities for in- and outbound are ensured by constraints (4) and (5). In combination with constraint (3) the capacity restriction is then ensured for the entire service. Constraints (6) ensure that the maximum number of available vehicles of one type in an external zone is never exceeded. Constraints (7) and (8) limit the number of urban vehicles at a satellite in each period in total and per transportation mode respectively. Finally, constraints (9) limit the maximum amount of demand which can be unloaded or loaded at a satellite in each period.

This reformulation has many similarities to known problems from the literature. Since we are only loading inbound demand and unloading outbound demand at an external zone, we are almost solving two bin packing problems (Dyckhoff and Finke, 1992) at the external zones, where the services are bins to which the demands are assigned to. Without constraint (3), the bin capacity constraints are ensured throughout the whole service if they are satisfied at the external zone.

This means also that for a known service schedule, the subproblem has several commonalities with a multiple knapsack problem with assignment restrictions (Dawande et al., 2000). The later accounting for the inbound and outbound conflicts and ensuring demand satisfaction.

The main difference with respect to the knapsack and the bin packing problem, which also makes our problem harder to solve, is the complex cost structure. Each service (bin) has a unique operating cost and each assignment of an item to a service has also a unique cost. Even complex bin packing formulations are not yet dealing with these kind of cost structures.

4 Solution method

To solve the problem efficiently, we use Benders decomposition (Benders, 1962). The idea is to decompose the problem into two easier solvable subproblems. Although the general methods is applied either to mixed-integer linear programs or stochastic programs, several authors adapted the method for integer programs (e.g., Laporte and Louveaux, 1993; Laporte et al., 2002; Restrepo et al., 2015). An overview of applications and acceleration techniques is given in Rahmaniani et al. (2016).

In Benders decomposition, the problem is decomposed into two subproblems. The master problem with the complicating variables is solved to generate lower bounds, while feasible solutions of the slave problems give upper bounds. In the following, we introduce first the general solution procedure and explain how to handle the integer subproblem. Then we define the slave problem in Section 4.2. The generated cuts from the slave problem are then used in the master problem (Section 4.3).

4.1 Benders Decomposition with Integer Decision Variables

In our problem, ρ_r are the complicating variables, which define the schedules. The easier variables are the demand assignment variables $x_{r^c,d,z}$. The master problem therefore selects the operated services, while in the slave problem a multiple knapsack problem with precedence constraints - due to the inbound-outbound loading - is solved. Our method benefits from the fact, that the linear relaxation of knapsack problems gives very good lower bounds. Therefore, we can generate the feasibility and the optimality cuts through the linear relaxation of the slave problem that are still tight. Moreover, when solving the master problem, we still derive lower bounds of our original integer problem.

In the slave problem, we first solve the dual formulation DSP to derive the classical Benders cuts. This solution gives an upper bound to the relaxed problem. If this upper bound is worse than the current best integer upper bound, also the integer problem will be worse and therefore the integer node is removed. If the relaxed upper bound is better, we solve the integer subproblem SP. This gives us a potential new upper bound. Because of the good linear relaxation of the integer slave problem, we can avoid solving too many subproblems. If the relaxed upper bound is better, the possibility of getting also a better true integer upper bound is high. To avoid the convergence to the relaxed problem, we further add the combinatorial cuts (21), which we describe in Section 4.3.

Algorithm 1 Benders decomposition

Set $\xi_{low} \leftarrow -\infty, \xi_{up} \leftarrow \infty$
while $\xi_{up} > \xi_{low}$ **do**
    solve $RMP \rightarrow \bar{\rho}_r, \xi_{low}$
    update core point
    solve $DSP(\bar{\rho}_r) \rightarrow \xi_{DSP}$
    if DSP is bounded **then**
        generate optimality cut (Magnanti-Wong)
        if $z_{DSP} < z_{up}$ **then**
            solve $SP(\bar{\rho}_r) \rightarrow \xi_{SP}$
            if $\xi_{SP} < \xi_{up}$ **then**
                $\xi_{up} \leftarrow \xi_{SP}$
            end if
        end if
    else
        generate feasibility cut
    end if
    generate combinatorial cut
end while

Algorithm 1 summarizes the general structure of the procedure. To further accelerate the solution method, the Benders decomposition algorithm is embedded into a Branch-and-Cut framework. At each integer node with a potential better solution, the subproblem is solved and cuts are generated.

4.2 Slave Problem

The slave problem $SP(\bar{\rho})$ for a given selection of services $\bar{\rho}$ is defined as follows:

$$\min \sum_{d\in\mathcal{D}} \sum_{r\in\mathcal{R}} \sum_{r^c\in\mathcal{R}^C(r)} \sum_{z\in\mathcal{Z}} \left(s_{d,z,t} + f_{d,e_r}\right) x_{r^c,d,z} \tag{12}$$

subject to

$$\sum_{z\in\mathcal{Z}} \sum_{r\in\mathcal{R}} \sum_{r^c\in\mathcal{R}^C(r)} x_{r^c,d,z} = 1 \qquad \forall d\in\mathcal{D} \tag{2}$$

$$x_{r^c,d_1,z_1} + x_{r^c,d_2,z_2} \leq 1 \qquad \forall r^c\in\mathcal{R}^C(r), r\in\mathcal{R}, d_1\in\mathcal{D}^I, d_2\in\mathcal{D}^O, z_1,z_2\in\mathcal{Z}(r), z_1\geq z_2 \tag{3}$$

$$\sum_{r\in\mathcal{R}(t,z)} \sum_{r^c\in\mathcal{R}^C(r)} \sum_{d\in\mathcal{D}} v_d x_{r^c,d,z} \leq u^V_{zt} \qquad \forall z\in\mathcal{Z}, t=1,\ldots,T \tag{9}$$

$$\sum_{z\in\mathcal{Z}} \sum_{d\in\mathcal{D}^I} v_d x_{r^c,d,z} \leq u^c_{\tau_r}\bar{\rho}_r \qquad \forall r^c\in\mathcal{R}^C(r), r\in\mathcal{R} \tag{13}$$

$$\sum_{z\in\mathcal{Z}} \sum_{d\in\mathcal{D}^O} v_d x_{r^c,d,z} \leq u^c_{\tau_r}\bar{\rho}_r \qquad \forall r^c\in\mathcal{R}^C(r), r\in\mathcal{R} \tag{14}$$

$$x_{r^c,d,z}\in\{0,1\} \qquad \forall d\in\mathcal{D}, r^c\in\mathcal{R}^C(r), r\in\mathcal{R}, z\in\mathcal{Z} \tag{15}$$

To generate optimality and feasibility cuts, the dual problem of the linear relaxation of $SP(\bar{\rho})$ is solved. We refer to that as $DSP(\bar{\rho})$. With $\alpha(d)$, $\beta(d_1,d_2,r^c,r,z_1,z_2)$, $\delta(z,t)$, $\gamma^{In}(r,r^c)$ and $\gamma^{Out}(r,r^c)$ being the dual variables of constraints (2),(3),(9), (13) and (14), the dual slave problem is defined as follows:

$$\max \sum_{d\in\mathcal{D}} \alpha(d) + \sum_{r\in\mathcal{R}} \sum_{r^c\in\mathcal{R}^C(r)} \sum_{d_1\in\mathcal{D}^I} \sum_{d_2\in\mathcal{D}^O} \sum_{z_1\in\mathcal{Z}(r)} \sum_{\substack{z_2\in\mathcal{Z}(r)\\ z_1\geq z_2}} \beta(d_1,d_2,r^c,r,z_1,z_2)$$
$$+ \sum_{r\in\mathcal{R}} \sum_{r^c\in\mathcal{R}^C(r)} u^c_\tau(r)\bar{\rho}_r \left(\gamma^{In}(r,r^c) + \gamma^{In}(r,r^c)\right) + \sum_{z\in\mathcal{Z}} \sum_{t=1}^{T} u^V_{zt}\delta(z,t) \tag{16}$$

subject to

$$\alpha(d) + \sum_{d_2 \in \mathcal{D}^O} \sum_{\substack{z_2 \in \mathcal{Z}(r) \\ z \geq z_2}} \beta(d, d_2, r^c, r, z, z_2) + v_d \left(\gamma^{In}(r, r^c) + \gamma^{In}(r, r^c)\right)$$

$$+ \sum_{\substack{t=1| \\ r \in \mathcal{R}(t,z)}}^{T} v_d \delta(z, t) \leq s_{d,z,t} + f_{d,e_r}$$

$$\forall d \in \mathcal{D}^I, r^c \in \mathcal{R}^C(r), r \in \mathcal{R}, z \in \mathcal{Z}(r) \tag{17}$$

$$\alpha(d) + \sum_{d_1 \in \mathcal{D}^I} \sum_{\substack{z_1 \in \mathcal{Z}(r) \\ z_1 \geq z}} \beta(d_1, d, r^c, r, z_1, z) + v_d \left(\gamma^{In}(r, r^c) + \gamma^{Out}(r, r^c)\right)$$

$$+ \sum_{\substack{t=1| \\ r \in \mathcal{R}(t,z)}}^{T} v_d \delta(z, t) \leq s_{d,z,t} + f_{d,e_r}$$

$$\forall d \in \mathcal{D}^O, r^c \in \mathcal{R}^C(r), r \in \mathcal{R}, z \in \mathcal{Z}(r) \tag{18}$$

$$\alpha \in \mathbb{R} \tag{19}$$

$$\beta, \gamma^{In}, \gamma^{Out}, \delta \leq 0 \tag{20}$$

Since the primal slave problem only consists of one decision variable, the dual problem has only one constraint. Because of the precedence constraint (3), we split the constraint into one for inbound and one for outbound demand. The resulting optimality cuts are improved according to Magnanti and Wong (1981) and the necessary core-points are updated according to Papadakos (2008).

4.3 Master Problem

Using the definitions of the slave problem, we can derive at each integer node either, if the dual slave problem is bounded, an optimality or otherwise a feasibility cut. We define the current set of optimality cuts as C^O and the current set of feasibility cuts as C^F, where c_o and c_f are a specific cut out of the optimality and the feasibility cut set. Moreover, we use the current set of combinatorial cuts C^C with c_c being one individual cut. For each integer node $\bar{\rho}$ with costs $\bar{\xi}_{\bar{\rho}}$ of the subproblem, the following combinatorial cut is derived:

$$\bar{\xi}_{\bar{\rho}} \left(\sum_{r \in \mathcal{R} | \bar{\rho}_r = 1} \rho_r - \sum_{r \in \mathcal{R} | \bar{\rho}_r = 0} \rho_r - \sum_{r \in \mathcal{R}} \bar{\rho}_r + 1 \right) \leq \xi \tag{21}$$

Similar to the idea of partial decomposition (Crainic et al., 2016a), we keep some information of the subproblem in the master problem. We use a row aggregation (Tsurkov,

2013) over the compartments and the fact that one demand is only assigned to one compartment. Thus, each demand is assigned to exactly one service. Moreover, only the relaxation is used, to avoid too many integer variables but still keeping enough information in the master problem.

Then, we define the current relaxed master problem $CRMP$ as follows:

$$\min \sum_{r \in \mathcal{R}} k_r \rho_r + \xi \tag{22}$$

subject to

$$\begin{aligned}
&(6), (7), (8) && \\
&c_o && \forall c_o \in C^O && (23)\\
&c_f && \forall c_f \in C^F && (24)\\
&c_c && \forall c_c \in C^C && (25)\\
&\sum_{d \in \mathcal{D}} \sum_{r \in \mathcal{R}} \sum_{z \in \mathcal{Z}} (s_{d,z,t} + f_{d,e_r})\, x_{rdz} \leq \xi && && (26)\\
&\sum_{z \in \mathcal{Z}} \sum_{r \in \mathcal{R}} x_{rdz} = 1 && \forall d \in \mathcal{D} && (27)\\
&\sum_{z \in \mathcal{Z}} \sum_{d \in \mathcal{D}^I} v_d x_{rdz} \leq n^c_{\tau_r} u^c_{\tau_r} \rho_r && \forall r \in \mathcal{R} && (28)\\
&\sum_{z \in \mathcal{Z}} \sum_{d \in \mathcal{D}^O} v_d x_{rdz} \leq n^c_{\tau_r} u^c_{\tau_r} \rho_r && \forall r \in \mathcal{R} && (29)\\
&\sum_{r \in \mathcal{R}(t,z)} \sum_{d \in \mathcal{D}} v_d x_{rdz} \leq u^V_{zt} && \forall t = 1, \dots, T, z \in \mathcal{Z} && (30)\\
&0 \leq x_{rdz} \leq 1 && \forall d \in \mathcal{D}, r \in \mathcal{R}, z \in \mathcal{Z} \xi \in \mathbb{R} && (31)\\
&\rho_r \in \{0, 1\} && \forall r \in \mathcal{R} && (32)
\end{aligned}$$

We minimize the service operating costs subject to the service related constraints from the IP formulation and the generated cuts. Moreover, (26) -(30) are the constraints related to the partial decomposition. The costs of the partial decomposition give a lower bound to the objective function through equation (26). In equations (28) and (29) are the capacity constraints. The capacity of the compartments is multiplied with the number of compartments to reflect the total capacity of the service. Equations (27) and (30) reflect again the demand assignment and the satellite capacity restrictions as before.

4.4 Valid Inequalities

To further strengthen the master problem relaxation, we introduce four sets of inequalities: two for inbound demands (33),(34) and two for outbound demands (35),(36).

$$\sum_{d \in \mathcal{D}^{I-}(t)} v_d \leq \sum_{r \in \mathcal{R}^{I-}(t)} u_{\tau_r} \rho_r, \qquad t = 1, \ldots, T \tag{33}$$

$$\sum_{d \in \mathcal{D}^{I+}(t)} v_d \leq \sum_{r \in \mathcal{R}^{I+}(t)} u_{\tau_r} \rho_r, \qquad t = 1, \ldots, T \tag{34}$$

$$\sum_{d \in \mathcal{D}^{O-}(t)} v_d \leq \sum_{r \in \mathcal{R}^{O-}(t)} u_{\tau_r} \rho_r, \qquad t = 1, \ldots, T \tag{35}$$

$$\sum_{d \in \mathcal{D}^{O+}(t)} v_d \leq \sum_{r \in \mathcal{R}^{O+}(t)} u_{\tau_r} \rho_r, \qquad t = 1, \ldots, T \tag{36}$$

$\mathcal{D}^{I-}(t) \subseteq \mathcal{D}$ defines the subset of inbound demands with a due date in period t or earlier ($\mathcal{D}^{I-}(t_1) \subseteq \mathcal{D}^{I-}(t_2)$ for $t_1 \leq t_2$). Since the due date of a demand may differ depending on the satellite which is used for the final distribution, the latest due date is used. Similar, $\mathcal{R}^{I-}(t) \subseteq \mathcal{R}$ are the services which can operate these inbound demands with deadline t. Then the volume of all demands until period t must be lower then the capacity of all operated services which can satisfy these demands. This is ensured by valid inequality (33). Inequality (34) considers instead of inbound due dates, the availability at external zones. $\mathcal{D}^{I+}(t) \subseteq \mathcal{D}$ are the inbound demands that are available in period t or later. Therefore, $\mathcal{D}^{I+}(t_1) \supseteq \mathcal{D}^{I+}(t_2)$ for $t_1 \leq t_2$ holds. This ensures that also in the last period enough capacity is available for resources which are only available in the last period.

Valid inequality (35) and (36) ensures the same for the outbound demand while here the due date at the external zone and the availability at satellites is considered.

5 Numerical Study

In this section, we first describe the numerical setup. Then, we show the performance of the proposed solution method in Section 5.2. For doing so, we compare our Benders decomposition algorithm to a commercial solver. In Section 5.3, we give insights on the benefits of using multi-modal transportation systems. The importance of considering both inbound and outbound flows in one network is shown in Section 5.4.

5.1 Numerical Setup

All experiments are performed on a cluster with Intel(R) Xeon(R) X5675 with 3.07GHz. Both the IP and the BD are implemented in C++ using Cplex Conert Technology 12.6.1.

To test and evaluate the performance of our method, we generate a set of new instances. We consider a planning horizon of 3 hours divided into 36 periods of 5 minutes. For the network, we generate 4 different network types with 2 external zones and 4, 6 and 2 times 8 satellites. In the default setting, each satellite has a capacity of 5000 and can dispatch one truck or one tram per period. For the distances we use the Euclidian distance between all points and different travel speeds depending on the vehicle are used. The networks are inspired by typical city structure but not based on real data. Illustrations of the 4 networks are shown in Appendix 6.

For the services, we generate tram services on predefined lines (which are also shown in the illustrations) and truck services. For the trams, we generate three times one tram line for each external zone and for one 8 satellite scenario only one tram line. The tram has 3 compartments with a capacity of 700 each. It visits a set of satellites and then returns via the same track in opposite direction and is operated several times in the planning horizon. For the trucks, a subset of satellites and a the external zone is selected. Moreover, the starting period of the service is generated randomly. For the basic scenario we assumed a large truck with capacity of 3000. When extending the network a smaller truck with 2000 capacity is further added. This resulted in 73, 71, 69 and 70 services. We assume that the tram does not have to deal with congestion and therefore is the faster transportation mode compared to the truck. Moreover, a fixed operating cost plus a total service distance depending costs is calculated and assigned to each service.

For the demand, we generate 24 demand instances with a demand of 150, 160, 170 and 180 and different shares of inbound and outbound demand between 50%, 67% and 83% of inbound traffic. The demand points are randomly distributed over the city center. Each demand is either an inbound or outbound demand and is assigned to a preferred external zone. Moreover, a volume between 50 and 100 and time windows for the availability and at the destination are generated. For the second tier, we define final distribution costs depending on the distance between the customer location and the location of the possible satellites.

5.2 Run Time Performance

First, we compare our Benders decomposition approach to the IP formulation solved by Cplex. We analyze the run time effect of the two main complexity drivers: the number of demand nodes and the number of services. For all instances a run time limit of 24 hours is used.

Table 1 shows the average run time of 24 instances each for different total demand scenarios. Each row consists of the 4 generated networks and the 6 generated demand scenarios. Cplex could only solve 78 of 96 instances within 24 hours to optimality while the Benders decomposition solved all instances. For the unsolved IPs, several instances

D	IP		BD		time improvement
	time (sec)	[solved]	time (sec)	[solved]	(in %)
150	13451	[22/24]	2100	[24/24]	84.39
160	19402	[20/24]	2430	[24/24]	87.48
170	29318	[18/24]	7297	[24/24]	75.11
180	35304	[18/24]	6459	[24/24]	81.70

Table 1: Run Time Depending on Number of Customers

had to be stopped because of memory problems and were resolved on a larger machine. For the instances which are not solved to optimality, the gap was between 0.15% and 0.01 %. Our approach, however, could solve all instances to optimality and further gives a run time improvement of more than 80% on average.

In a second study, we analyze the effect of increasing the number of services. For that we use the six 150 demand scenarios and the the four generated networks. Then, we added 10, 20, and 30 services to the existing ones.

R	IP		BD		time improvement
	time (sec)	[solved]	time (sec)	[solved]	(in %)
basic	11455	[22/24]	1829	[24/24]	84.03
basic +10	8296	[22/24]	2061	[24/24]	75.16
basic +20	20770	[20/24]	2753	[24/24]	86.74
basic +30	33609	[16/24]	4383	[24/24]	86.96

Table 2: Run Time Depending on Number of Services

Table 2 shows that Cplex could only solve 80 of the 96 instances within 24 hours to optimality. Again, due to memory problems several IP instances had to be stopped and were resolved on a larger machine. For the instances which are not solved to optimality, the gap was between 0.22% and 0.01%. Especially when increasing the number of services, the number of solvable instances decreases. For the Benders decomposition, we see the classical behavior that the run time is increasing with the number of master variables since the complexity in the master problem increases. Overall we can state that our approach clearly outperforms Cplex.

To further test the limits of our methods, we increase the number of demands and the number of services. First, we increase the demand scenarios up to 250 customers. Table 3 shows that for 200 demand nodes, we still solve 22 out of 24 instances to optimality and the unsolved instances end with an average gap below 3%. For 250 demand nodes, half of the instances are still solved to optimality and the other half with an average gap of 2%.

For the services, we double the number of services of the basic scenario, resulting with

D	time (sec)	[solved]	gap of unsolved (in %)
200	16023	[22/24]	2.97
220	29070	[19/24]	2.04
250	42855	[12/24]	1.99

Table 3: Run Time BD Depending on Number of Demands (larger instances)

approximately 140 services. Then, we further added 20 services to the network.

R	time (sec)	[solved]	gap of unsolved (in %)
~ 140	8044	[24/24]	-
~ 160	10508	[23/24]	1.44

Table 4: Run Time BD Depending on Number of Services (larger instances)

Table 4 shows that even for the largest scenario 23 out of 24 instances are solved to optimality and the unsolved scenario ends with a gap of 1.44%.

Summarized, this shows that our method does not only outperform Cplex, even for larger instances many scenarios are solved to optimality or the gap is small.

5.3 Benefits of a Multi-Modal Network

In this section, we use the proposed method to analyze the benefits of the different transportation modes and the effect of different inbound and outbound shares. We use the 150 customers scenarios with inbound and outbound shares of 50%, 66% and 83%. For the network, we use the basic scenario of the 4 satellite network of the previous section and considered either both modes (truck and tram) or only one.

	total costs	# services	start util. (%)	end util. (%)	empty km	full (80%) km
tram + truck	22624	5.7 [3.7/2.0]	54 [66/44]	28 [38/18]	125 [13/112]	122 [104/18]
tram + truck	22624	5.7 [3.7/2.0]	54 [66/44]	28 [38/18]	125 [13/112]	122 [104/18]
tram	23464	6.3	56	29	39	109
truck	24457	6.2	41	21	328	32

Table 5: Analysis of the Fleet Structure

Table 5 compares the mixed fleet to the two scenarios with single mode fleet structures. Besides the total costs and the number of services, the utilization of the vehicle at the start and end external zone, and the full and empty truck kilometers are shown. Hereby, we assume that a truck is full if more than 80% of its capacity is used in a segment.

For the multi-modal fleet case, the results are further split up into tram and truck in brackets.

One can see that the multi-modal fleeet structure significantly reduces the costs and also the number of operated services. Due to multiple compartments which allow combining inbound and outbound demands, the tram is almost never empty. This is not only true for the tram scenario but also for the multi-modal scenario. It is clear that the utilization at the start external zone is higher than at the end external zone, since more demand is entering the city than demand is leaving.

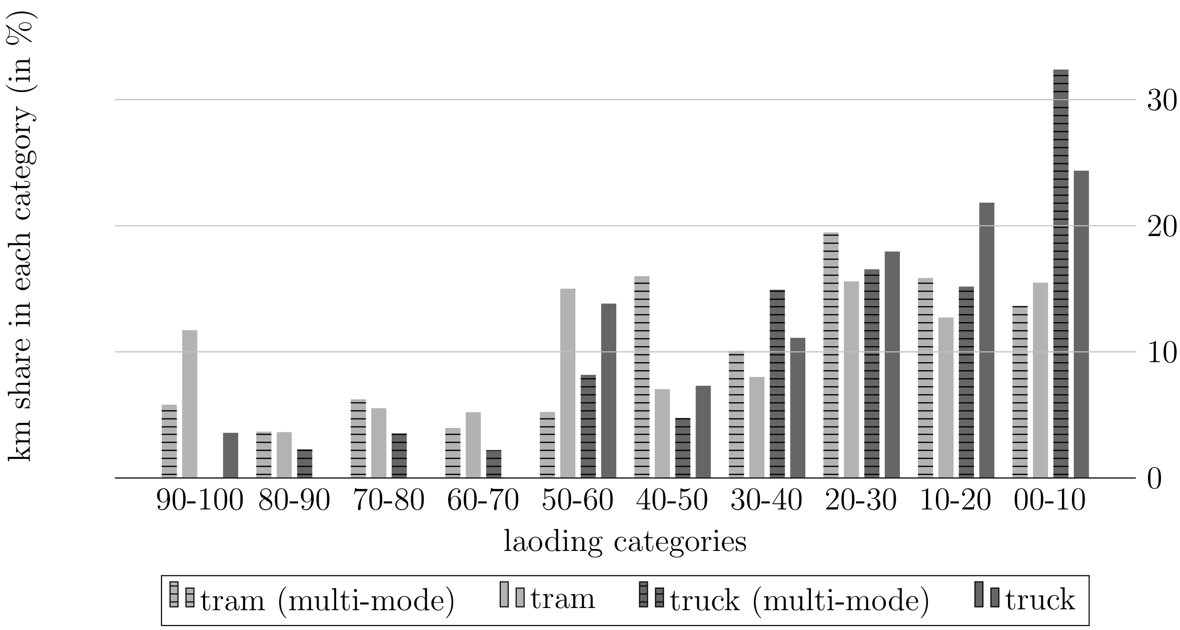

Figure 3: Load distribution

In Figure 3, we further analyze the loading of the vehicle. Therefore, we categorized the loading level into the 10 different percentage categories. In each bar, we show the percentage of the total distance driven in that category. In both single mode scenarios relatively more kilometers are driven with a utilization of more than 90% and also with 50% to 60% loading. Interestingly, in the truck scenario, no truck is driving with a load between 60% and 90%. This means that the trucks are loaded especially to deliver at the first satellite and then continue the service. They have the need to completely unload the truck fast, to be able to load again goods for the outbound traffic at a new satellite. Trams on the other side can benefit from the multiple compartments. They just need to empty one compartment and can therefore also use the satellites for inbound and outbound demand in parallel. Nevertheless, a single tram system is also not preferable. Since the trams drive back and forth on the line, the travel time is typically slower than the truck. The reduction of services, when combining both modes, shows that the model is taking advantages of the strengths of both modes.

5.4 Benefits of Combining Inbound and Outbound Demands

In this section, we show the benefits of combining both inbound and outbound demands in one network. We use the same setup as in the previous section. However, this time we compare the results of our model with the results of solving a inbound demand model and a outbound demand model separately. This means that a service can not be used for both demand types, even if the compartment is empty. Table 6 shows the improvement when combining both demands in one model compared to scheduling them separately. Since it is obvious that for a pure inbound demand model the truck utilization at the end of the service is zero, we only compare the utilization of inbound services with each other. The same applies for outbound services.

$\|\mathcal{D}^I\|/\|\mathcal{D}\|$	scenario	total costs	# services	start util.	end util.	empty km	full km (80 %)
0.5	tram + truck	-15.61	-33.33	6.19	-45.69	-82.57	-63.87
	tram	-24.17	-33.33	-14.98	-50.00	-100.00	-78.20
	truck	-12.21	-31.25	0.00	-55.91	-62.37	-74.84
0.67	tram + truck	-12.73	-25.00	6.34	-56.95	-84.72	-14.92
	tram	-17.14	-27.78	0.00	-62.93	-86.73	-50.60
	truck	-8.23	-22.22	0.00	-71.43	-52.19	-100.00
0.83	tram + truck	-13.08	-25.00	0.00	-61.50	-71.44	-5.66
	tram	-16.00	-27.78	0.00	-63.16	-93.28	-18.07
	truck	-10.27	-25.00	0.00	-66.67	-42.88	-41.89

Table 6: Improvement when Combining Inbound and Outbound Demands (in %)

The results show that the costs and the number of services are reduced. Moreover, this effect is clearly the strongest for the demand scenario with 50% inbound and 50% outbound demands, as both flows are balanced out. While also the number of empty kilometers is reduced, surprisingly also the number of full kilometer is reduced. This, however, is due to the fact that for combining inbound and outbound flows, the compartment or truck needs to be empty at a certain point of time and therefore, it should not be loaded too full.

Table 6 further indicates that the benefits are the strongest in the pure tram scenario and the weakest in the truck scenario. This shows that the multiple compartment structure is used to start already with outbound demand while some compartments are still used for inbound demands.

The results of the truck utilization are no surprisingly. While for the utilization for inbound demand at the start of the service is very similar, for the outbound demand, the utilization is worse in the combined scenario. The outbound demand is assigned to the available services which delivered already inbound demand. This, however, means that it is not always the best choice. In a pure outbound scenario, less services are used for the outbound demand. And this results in a higher utilization. However, looking at the whole system, this is not an achievable goal. These effects are also in line with the load distribution analysis of the previous section.

Summarized, we can state that the combination of both in one model should be considered to efficiently use the resources. Moreover, the utilization and the number of full and empty kilometers has to be treated carefully as performance measures in such a system.

6 Conclusion and Outlook

In this paper, we introduced a new model for a 2T-CL system which considers both inbound and outbound traffic and multi-modal services. To solve the problem, we introduced a new formulation of the problem and presented a Benders decomposition algorithm for integer problems. The performance is improved by an implementation within a Brach-and-Cut framework and by valid inequalities, partial decomposition techniques, and pareto-optimal cuts. Extensive numerical studies have shown that the solution method outperforms Cplex and that the consideration of multi-modal services gives a city logistics network more flexibility and that inbound and outbound flows should be considered in one optimization model.

While we are using the proposed formulation to develop optimal solution method based on Benders decomposition, it opens further directions for large scale solution methods. Mat- or meta-heuristics can especially benefit from the well-studied knapsack subproblem.

Acknowledgments

We gratefully acknowledge the financial support provided by the Natural Sciences and Engineering Council of Canada (NSERC), through its Discovery and Acceleration Grant programs. We also gratefully acknowledge the support of Fonds de recherche du Québec through their Strategic Clustering Grant program, the Canadian Foundation for Innovation and the Ministère de l'Économie, de la Science et de l'Innovation of Quebec through the Leaders Fond Infrastructure Grant program, as well as Calcul Québec and Compute Canada for access to their high-performance computing infrastructure. While working on this project, the second author was also Adjunct Professor, Department of Computer Science and Operations Research, Université de Montréal.

References

Andersen, J., Crainic, T. G., and Christiansen, M. (2009a). Service network design with asset management: Formulations and comparative analyses. *Transportation Research Part C: Emerging Technologies*, 17(2):197–207.

Andersen, J., Crainic, T. G., and Christiansen, M. (2009b). Service network design with management and coordination of multiple fleets. *European Journal of Operational Research*, 193(2):377–389.

Bektas, T., Crainic, T. G., and Van Woensel, T. (2015). From managing urban freight to smart city logistics networks. Publication CIRRELT-2015-17, Centre interuniversitaire de recherche sur les réseaux d'entreprise, la logistique et le transport, Université de Montréal, Montréal, QC, Canada.

Benders, J. F. (1962). Partitioning procedures for solving mixed-variables programming problems. *Numerische mathematik*, 4(1):238–252.

Crainic, T. G. (2000). Service network design in freight transportation. *European Journal of Operational Research*, 122(2):272–288.

Crainic, T. G., Errico, F., Rei, W., and Ricciardi, N. (2012). Integrating c2e and c2c traffic into city logistics planning. *Procedia-social and behavioral sciences*, 39:47–60.

Crainic, T. G., Errico, F., Rei, W., and Ricciardi, N. (2015a). Modeling demand uncertainty in two-tier city logistics tactical planning. *Transportation Science*, 50(2):559–578.

Crainic, T. G., Hewitt, M., Maggioni, F., and Rei, W. (2016a). Partial benders decomposition strategies for two-stage stochastic integer programs. Publication CIRRELT-2016-37, Centre interuniversitaire de recherche sur les réseaux d'entreprise, la logistique et le transport, Université de Montréal, Montréal, QC, Canada.

Crainic, T. G., Hewitt, M., Toulouse, M., and Vu, D. M. (2014). Service network design with resource constraints. *Transportation Science*.

Crainic, T. G., Hewitt, M., Toulouse, M., and Vu, D. M. (2015b). Location and service network design. Publication CIRRELT-2015-45, Centre interuniversitaire de recherche sur les réseaux d'entreprise, la logistique et le transport, Université de Montréal, Montréal, QC, Canada.

Crainic, T. G., Hewitt, M., Toulouse, M., and Vu, D. M. (2016b). Scheduled service network design with resource acquisition and management. Publication CIRRELT-2016-50, Centre interuniversitaire de recherche sur les réseaux d'entreprise, la logistique et le transport, Université de Montréal, Montréal, QC, Canada.

Crainic, T. G. and Montreuil, B. (2016). Physical internet enabled hyperconnected city logistics. *Transportation Research Procedia*, 12:383–398.

Crainic, T. G., Ricciardi, N., and Storchi, G. (2004). Advanced freight transportation systems for congested urban areas. *Transportation Research Part C: Emerging Technologies*, 12(2):119–137.

Crainic, T. G., Ricciardi, N., and Storchi, G. (2009). Models for evaluating and planning city logistics systems. *Transportation science*, 43(4):432–454.

Crainic, T. G. and Sgalambro, A. (2014). Service network design models for two-tier city logistics. *Optimization Letters*, 8(4):1375–1387.

Dawande, M., Kalagnanam, J., Keskinocak, P., Salman, F. S., and Ravi, R. (2000). Approximation algorithms for the multiple knapsack problem with assignment restrictions. *Journal of Combinatorial Optimization*, 4(2):171–186.

Dyckhoff, H. and Finke, U. (1992). *Cutting and packing in production and distribution: A typology and bibliography*. Springer Science & Business Media.

Ghilas, V., Demir, E., and Van Woensel, T. (2013). Integrating passenger and freight transportation: model formulation and insights. In *Beta Working Papers WP 441*. Technische Universiteit Eindhoven.

Hemmelmayr, V. C., Cordeau, J.-F., and Crainic, T. G. (2012). An adaptive large neighborhood search heuristic for two-echelon vehicle routing problems arising in city logistics. *Computers & operations research*, 39(12):3215–3228.

Laporte, G. and Louveaux, F. V. (1993). The integer l-shaped method for stochastic integer programs with complete recourse. *Operations research letters*, 13(3):133–142.

Laporte, G., Louveaux, F. V., and Van Hamme, L. (2002). An integer l-shaped algorithm for the capacitated vehicle routing problem with stochastic demands. *Operations Research*, 50(3):415–423.

Lindholm, M. and Behrends, S. (2012). Challenges in urban freight transport planning–a review in the baltic sea region. *Journal of Transport Geography*, 22:129–136.

Magnanti, T. L. and Wong, R. T. (1981). Accelerating benders decomposition: Algorithmic enhancement and model selection criteria. *Operations Research*, 29(3):464–484.

Neumann-Saavedra, B. A., Crainic, T. G., Gendron, B., Mattfeld, D. C., and Römer, M. (2016). Service network design of bike sharing systems with resource constraints. In *International Conference on Computational Logistics*, pages 352–366. Springer.

Papadakos, N. (2008). Practical enhancements to the magnanti–wong method. *Operations Research Letters*, 36(4):444–449.

Perboli, G., Tadei, R., and Vigo, D. (2011). The two-echelon capacitated vehicle routing problem: models and math-based heuristics. *Transportation Science*, 45(3):364–380.

Rahmaniani, R., Crainic, T. G., Gendreau, M., and Rei, W. (2016). The Benders decomposition algorithm: A literature review. *European Journal of Operational Research*.

Restrepo, M. I., Gendron, B., and Rousseau, L.-M. (2015). Combining benders decomposition and column generation for multi-activity tour scheduling. Publication CIRRELT-2015-57, Centre interuniversitaire de recherche sur les réseaux d'entreprise, la logistique et le transport, Université de Montréal, Montréal, QC, Canada.

Savelsbergh, M. and Van Woensel, T. (2016). City logistics: Challenges and opportunities. *Transportation Science*, 50(2):579–590.

Tsurkov, V. (2013). *Large-scale Optimization: Problems and Methods*, volume 51. Springer Science & Business Media.

United Nations (2014). World urbanization prospects, the 2014 revision. Technical report, United Nations, Department of Economic and Social Affairs, Population Division,New York.

Wang, X., Crainic, T. G., and Wallace, S. W. (2016). Stochastic scheduled service network design: The value of deterministic solutions. Publication CIRRELT-2016-14, Centre interuniversitaire de recherche sur les réseaux d'entreprise, la logistique et le transport, Université de Montréal, Montréal, QC, Canada.

Appendix 1 - Notation

Notation	Description
$t = 1, \ldots, T$	planning horizon
$\mathcal{M}$	set of means of transport
$\mathcal{E}$ ($\mathcal{E}^m$)	set of external zones (of mode m)
$\mathcal{Z}$ ($\mathcal{Z}^m$)	set of satellites (of mode m)
$\mathcal{Z}(d)$	set of feasible satellites for demand d
$\mathcal{T}$ ($\mathcal{T}^m$)	set of urban-vehicle types (of mode m)
$\mathcal{D}$	set of demands
$\mathcal{D}^I$, $\mathcal{D}^O$	set of inbound, outbound demands
$\mathcal{R}$	set of urban-vehicle services
$\mathcal{R}^C(r)$	set of compartment services of urban-vehicle service r
$n_{e\tau}$	fleet size of urban-vehicle τ at external zone e
u_τ (u^c_τ)	(compartment) capacity of urban-vehicle τ
n^c_τ	number of compartments in urban-vehicle τ
$u^{\mathcal{T}}_{zt}$ (u^m_{zt})	urban-vehicle limit at satellite z in period t (for mode m)
u^V_{zt}	loaded/unloaded volume limit at satellite z in period t
v_d	volume of demand d
$g_{ij}(t)$	travel time from i to j in period t
h_τ	service time of urban vehicle τ
w_{zr}	waiting time of service r at satellite z
e_r	external zone of urban-vehicle service r
τ_r	urban-vehicle type of urban-vehicle service r
m_r	mode of of urban-vehicle service r
σ^r	ordered set of visited satellites of urban-vehicle service r
t^r_0, t^r_i	departure time of urban-vehicle service r at the external zone, the ith satellite
k_r	operating costs of urban-vehicle service r
f_{de}	costs for assigning demand d to external zone e
s_{dzt}	costs for assigning demand d to satellite z in period t

Table 7: Notation

Appendix 1 - Networks

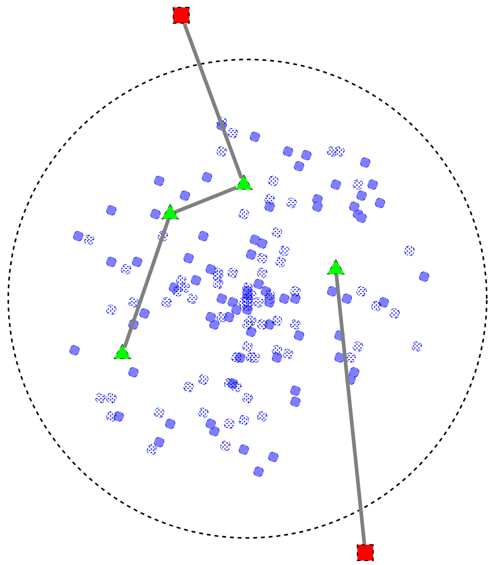

Figure 4: Network 1

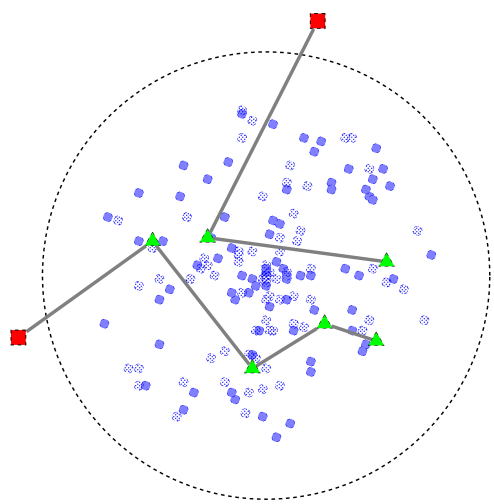

Figure 5: Network 2

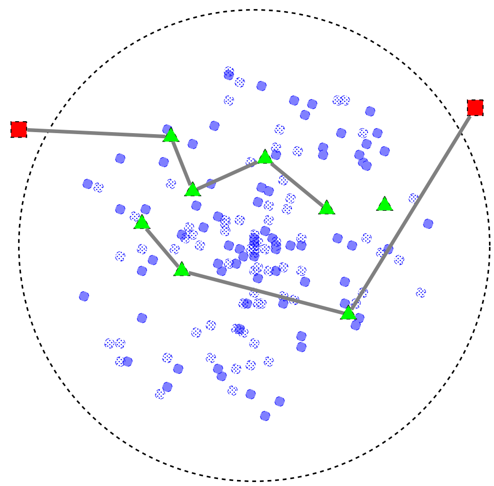

Figure 6: Network 3

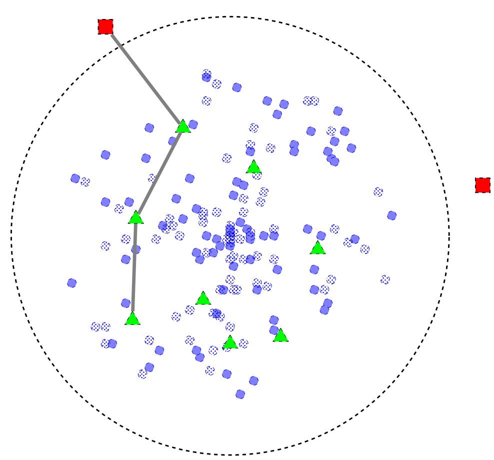

Figure 7: Network 4