CIRRELT

CIRRELT-2024-22

Resource Planning and Cost Allocation for Tactical Planning in Cooperative Two-Tier City Logistics Systems

Johannes Gückel
Teodor Gabriel Crainic
Pirmin Fontaine

July 2024

Bureau de Montréal
Université de Montréal
C.P. 6128, succ. Centre-Ville
Montréal (Québec) H3C 3J7
Tél : 1-514-343-7575
Télécopie : 1-514-343-7121

Bureau de Québec
Université Laval,
2325, rue de la Terrasse
Pavillon Palasis-Prince, local 2415
Québec (Québec) G1V 0A6
Tél : 1-418-656-2073
Télécopie : 1-418-656-2624

Resource Planning and Cost Allocation for Tactical Planning in Cooperative Two-Tier City Logistics Systems

Johannes Gückel[1], Teodor Gabriel Crainic[2,*], Pirmin Fontaine[1]

1. Ingolstadt School of Management, Mathematical Institute for Machine Learning and Data Science, Catholic University of Eichstätt-Ingolstadt

2. Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT) and School of Management, Université du Québec à Montréal

Abstract. The rapid transport of freight is an essential feature of modern societies and an enabling factor for economic trade and growth. Nevertheless, the negative impact of freight transportation in urban areas poses challenges for Logistics Service Providers (LSPs) as well as for municipalities. In this context, a centrally coordinated Two-Tier City Logistics System (2T-CLS), in which LSPs voluntarily agree to cooperate with each other, has the potential to reduce both economic and environmental impact costs. In order to plan such a system, it is important not only to make efficient use of the resources provided but also to have a mechanism that allocates the costs incurred to the individual LSPs. We introduce a mixed-integer linear program (MILP) formulation for the tactical planning of a 2T-CLS involving multiple LSPs that share their resources and customer demands. This MILP comprises a service network design formulation on the first tier and a vehicle routing problem formulation on the second tier, which are connected with each other. To address larger instances, we introduce an Integrative Two-Step Large Neighborhood Search with adaptive components that integrates first and second-tier decisions. In order not only to minimize the costs incurred but also to distribute them fairly, we investigate different problem-specific proportional methods, as well as more advanced game theoretical methods. Numerical experiments show that cooperation leads to average cost savings of 26.91%, which primarily stems from first-tier cooperation.

Keywords: Logistics, City Logistics, large neighborhood search, horizontal cooperation, cost allocation.

Acknowledgements. The paper was written as part of a project funded by the German Research Foundation. We would like to express our gratitude to the German Research Foundation (project number: ROCOCO FO1468/1-1) for the financial support of our research project. Additionally, our gratitude goes to the Bavarian Research Alliance for their support of the research stay in Montreal, which contributed significantly to the development of this paper. While working on this paper, the second author was Adjunct Professor, Department of Computer Science and Operations Research, Université de Montréal, Canada. We gratefully acknowledge the financial support provided by the Canadian Natural Sciences and Engineering Research Council (NSERC) through its Discovery grant program, as well as by the Fonds de recherche du Québec (FRQ) through their CIRRELT infrastructure grants.

---

* Corresponding author: teodorgabriel.crainic@cirrelt.net

Dépôt légal – Bibliothèque et Archives nationales du Québec
Bibliothèque et Archives Canada, 2024

1 Introduction

The transportation of goods and people is an indispensable part of modern societies and economies. At the same time, transportation is a major disruptive factor, especially for urban life, due to congestion, noise, emissions, space consumption, and other negative external effects. According to a United Nations (2018) forecast, urbanization will increase from 55% in 2018 to 68% in 2050. Simultaneously, the increasing degree of digitalization is leading to a trend towards more online orders, which means that e-commerce volume will rise as well (Lone et al., 2021). These two developments result in an increasing number of goods and people that need to be transported within cities. Making cities sustainable and environmentally friendly calls for urban logistics concepts that meet this increasing demand as efficiently as possible. While the need for efficient and sustainable transport concepts for inner-city transportation is more important than ever, Logistics Service Providers (LSPs) struggle with low load factors, empty trips, long dwell times at loading and unloading points, and a large number of deliveries to individual customers (Cepolina and Farina, 2015).

The goal of city logistics is to efficiently manage the transportation of goods while minimizing the negative consequences of transportation (Savelsbergh and van Woensel, 2016). This can be achieved either by a single LSP operating in a city as it is in Monaco or by cooperation, i.e., resource and/or demand sharing among the operating LSPs (Fontaine et al., 2023). While the first option might be difficult to implement in many cities for political reasons, cooperation receives resistance from LSPs because they still want customer contact. This raises the question of whether cooperation on both tiers is necessary and how much the benefits are. While cooperations in the area of vehicle routing in various single-level environments have already received attention in the literature, cooperation in the area of Two-Tier City Logistics Systems (2T-CLSs) is a rather unexplored topic. In a survey, Cruijssen et al. (2007a) pointed out that LSPs hold a strong belief in the potential efficiency gains resulting from cooperation. Further, Cruijssen et al. (2007b) conducted a survey on a considerable number of LSPs in Belgium and found that despite the clear advantages of cooperation, developing a justifiable cost-sharing plan poses a significant obstacle to LSPs' joint ventures.

In order to maintain the stability of cooperation among partners, the implementation of a cost allocation mechanism that is perceived as fair by all LSPs is essential. Such a mechanism would provide each partner in the coalition with a strong incentive to participate within the coalition. Previous studies on the cooperation of LSPs assumed cooperation in single tier settings only. This paper specifically focuses on cooperation among LSPs in the tactical planning of 2T-CLSs. This enables us to consider the impacts of cooperation on both tiers, both individually and jointly, and thus also to evaluate interaction effects. Introduced by Crainic et al. (2004), these systems feature the preliminary delivery of freight to City Distribution Centers (CDCs) located on the outskirts of urban areas. From these CDCs, large urban vehicles deliver freight to satellites — small facilities located in the inner city. Proceeding from the satellites, the last-mile deliveries to the customer location take place using vehicles, hereinafter referred to as city freighters, that are smaller, environmentally friendly, and cost-efficient. This two-tier structure leads to challenging optimization problems because of the appearance of NP-hard problems on both tiers that require connection and synchronization with each other.

Until today it is still unclear how the resources and customers are deployed and shared among the LSPs in a tactical plan since algorithmic solutions are still limited. Furthermore, it is unclear how the costs arising from the tactical plan are allocated to the individual LSPs. This paper fills this research gaps by making four key contributions.

1. We introduce an innovative Mixed-Integer Linear Program (MILP) formulation for tactical planning of a 2T-CLS with cooperating LSPs that combines and connects both, a service network design formulation on the first tier and a vehicle routing problem formulation on the second tier.
2. We develop a problem-specific two-step large neighborhood search with adaptive components that exploits the two-tier problem structure with problem-specific operators and is able to solve larger instances efficiently.
3. We introduce problem-specific proportional methods and apply them, along with game theoretical methods, to allocate costs to the LSPs.
4. Through an extensive numerical study, we prove the performance of our metaheuristic and provide valuable managerial insights on resource and demand sharing, as well as on cost allocation in 2T-CLSs that is of high relevance for LSPs as well as for municipalities.

The paper is organized as follows: After the literature review in Section 2, we detail the problem setting in Section 3. We present the MILP for resource planning in Section 4. Section 5 details our metaheuristic. Section 6 introduces problem-specific proportional methods and presents game theoretical cost allocation methods. In Section 7, we conduct an extensive numerical study regarding the performance of the metaheuristic as well as cost benefits and cost allocations. The paper concludes with a summary and future research directions in Section 8.

2 Literature review

The literature review is divided into four sections. Section 2.1 presents literature related to 2T-CLSs. Section 2.2 focuses on cooperation in transportation, while Section 2.3 reviews the literature on cost allocation methods. Finally, Section 2.4 concludes the literature review and outlines the research gap.

2.1 Two-tier city logistics

Consideration of 2T-CLSs in the scientific community was introduced by the work of Crainic et al. (2004), which addressed the problem of the optimal location of satellites. A case study with data from the city of Rome clearly showed that a 2T-CLS can greatly reduce the distance traveled by large trucks within the city and, in return, smaller, more environmentally friendly vehicles can be used for the final delivery to the customers. The general modeling framework for the tactical planning problem in a 2T-CLS was introduced by Crainic et al. (2009). They took into account the time dimension of the demand and the associated need to synchronize and schedule the vehicles of the first- and second tier. Nevertheless, only inbound demand was considered. In Crainic and Sgalambro (2014), a service network design formulation was proposed for the tactical planning problem in a 2T-CLS, including a discussion on algorithmic solution perspectives. Crainic et al. (2016) explicitly considered demand uncertainty in tactical planning. They proposed a two-stage stochastic programming formulation and presented different strategies for adjusting the plan to the observed demand. Crainic et al. (2021) proposed a scheduled service network design formulation as a modeling framework for tactical planning of a 2T-CLS and adapted it to the specific problem characteristics. Fontaine et al. (2021) expanded the existing literature with more realistic assumptions and considered both inbound and outbound demand as well as different modes of transportation. They presented a scheduled service network design formulation and developed an efficient Benders decomposition algorithm to solve it. In their investigation, they demonstrated that a multi-modal fleet on the first tier of a two-tier system can reduce costs and improve utilization. Considering the decisions about the number and location of facilities on both tiers, Winkenbach et al. (2016) presented a MILP for the two-echelon capacitated location routing problem. Schmidt et al. (2022) investigated the integration of public transport service providers in a two-tier system to deliver freight from the outskirts to satellite depots. Further Fontaine et al. (2023) recently investigated under which circumstances a single-tier or a two-tier system is beneficial to a city. Related to the literature on 2T-CLSs, Perboli et al. (2021) addressed the joint problem of operating satellites and providing services to customers.

The only investigation explicitly considering different LSPs cooperating in a 2T-CLS was conducted by Crainic et al. (2020). They assumed that demand could be satisfied by any participant of the coalition of LSPs. As they pointed out, the cooperation leads to significant efficiency improvements in terms of monetary costs and environmental footprint. However, they did not consider the allocation of costs to the LSPs. On the operational level, many publications exist regarding two-echelon vehicle routing problems. For a more extensive investigation into this area, we refer to the literature review of Sluijk et al. (2023).

2.2 Cooperation in transportation

Cooperation in transportation can take on different forms, including horizontal and vertical cooperation. Vertical cooperation involves collaboration between organizations at different levels of the supply chain, such as suppliers, retailers, and LSPs. On the other hand, horizontal cooperation refers to collaboration among organizations operating at the same level of the supply chain. Furthermore, diagonal cooperation describes models that incorporate both horizontal and vertical cooperation (see Rusich et al. (2017) for a framework of collaborative logistics). In this work, we focus on horizontal cooperation. According to the EU Commission (2001), horizontal cooperation refers to the collaboration through an agreement or a concerted practice between companies operating at the same level in the market that are potentially competing with each other. While horizontal cooperation in the maritime and airline industries has been used and explored for some time in urban freight transport, it is still a growing area of research.

The literature on horizontal cooperation consists largely of studies on the classification of forms of cooperation and empirical studies on the possible benefits and barriers of horizontal cooperation. According to Cruijssen et al. (2007b), horizontal cooperation can be used to identify and exploit win-win situations between different companies at the same level in the supply chain. The cooperation can be designed in different ways. As Pérez-Bernabeu et al. (2015) pointed out, the cooperating companies can be competing or unrelated suppliers, manufacturers, retailers, receivers, or LSPs that share information, facilities, or resources with the goal of reducing costs and/or improving service.

In the context of transportation logistics, horizontal cooperation has been considered to be an efficient instrument for the sustainable and efficient design of freight transport and has gained increased attention in recent years (Pan et al., 2019). Agarwal and Ergun (2010) proposed a mechanism aimed at directing liner shipping carriers to prioritize the collective interests of the alliance while also optimizing their own profitability. Nataraj et al. (2019) investigate the problem of jointly opening a CDC and conclude that significant overall cost savings can be achieved with a higher intensity of cooperation. Regarding cooperation at satellite depots, Bruni et al. (2024) explored the advantages of vertical and horizontal cooperation, identifying considerable cost reductions for both types. For a literature review on problems, approaches and further research areas related to the horizontal cooperation of LSPs, we refer to Pan et al. (2019).

In recent years, there have been a number of publications that have presented the potential benefits of cooperative planning in single-tier settings in terms of total profit improvements (e.g., Montoya-Torres et al., 2016). The potential profit improvements achieved in the respective papers through cooperation ranged between 20% and 30% (Gansterer and Hartl, 2018). Nonetheless, when constructing stable cooperations, the fair treatment of each coalition member is crucial. In order to ensure a certain degree of fairness, constraints addressing the workload of LSPs are used in some studies (e.g., Mancini et al., 2021; Gansterer et al., 2018). Recently, several papers used auction theory for operational planning in collaborative vehicle routing problems (Gansterer et al., 2020). For a comprehensive review of collaborative vehicle routing, we refer to Gansterer and Hartl (2018). In their review on city logistics, Parisa et al. (2019) pointed out that while collaborative approaches are extensively studied in theory, their implementation in practice is difficult due to the lack of a business model. This underlines the urgent need for a fair cost allocation mechanism to distribute the benefits of the cooperation to the participating LSPs.

2.3 Cost allocation methods

Cost allocation in cooperative games has already been studied in the literature for a variety of use cases. For a detailed review on cost allocation mechanisms in collaborative transportation, we refer to Guajardo and Rönnqvist (2016) in which over 40 different allocation mechanisms were identified in the reviewed literature. The mechanisms most commonly used in practice are based on proportional methods, where each participant is allocated a cost share in relation to a predefined reference value. In the simplest case, each participant bears an equal share of the costs. Other classic criteria are, for example, the share of demand or the stand-alone costs (Guajardo and Rönnqvist, 2016). Widely used methods based on cooperative game theory are the Shapley Value (Shapley, 1953), the core allocation (Gillies, 1959), and the nucleolus (Schmeidler, 1969). In addition, cost allocation mechanisms were introduced in the context of specific applications, such as the Equal Profit Method (EPM) developed by Frisk et al. (2010) in the context of collaborative forest transportation. Further cost allocation mechanisms based on the distinction between separable and non-separable costs are introduced by Tijs and Driessen (1986).

Vanovermeire et al. (2014) compared different cost allocation methods in collaborative transport, especially with partners who have different characteristics. They showed that the choice of the appropriate cost allocation method depends heavily on the characteristics of the coalition. An example of cost allocation in horizontal carrier cooperation can be found in Verdonck et al. (2016), where they investigated the cooperative carrier facility location problem, comparing three different cost allocation mechanisms. Kimms and Kozeletskyi (2016b) determined the Shapley Value for the cooperative traveling salesman problem. Further, Kimms and Kozeletskyi (2016a) proposed an a priori core cost allocation for a horizontal cooperating traveling salesman, which provided expected costs for the coalition participants.

2.4 Research gap

Overall, it can be stated that cooperative aspects in different single-tier settings have already been considered in the literature. In contrast, cooperative aspects in 2T-CLSs have hardly received any attention, although it is particularly necessary in this context, as several LSPs operate within the same system. Until today, there is no efficient solution method for tactical planning in a multi-modal 2T-CLS with multiple LSPs combining service network design on the first tier and vehicle routing on the second tier. Although various cost allocation mechanisms in collaborative transportation have been discussed in the literature, there is no publication to date that examines cost allocation in 2T-CLSs.

3 Problem setting

We describe the general functioning of 2T-CLSs in Section 3.1. Section 3.2 describes the tactical planning in 2T-CLSs. Section 3.3 outlines the cooperative aspects of the problem.

3.1 2T-CLS

For the illustration of 2T-CLSs, we use the terms introduced by Crainic et al. (2009). A 2T-CLS consists of two levels of physical infrastructure, the CDCs and the satellite platforms.

CDCs represent the first tier of facilities and are typically located on the outskirts of a city. Freight is first delivered from its origin location to the CDCs. For this, we assume costs for the delivery of each customer demand to each CDC. At the CDCs, the loads are sorted and consolidated for further distribution to satellites via urban vehicle services. A service starts at a certain time period at a CDC, visits one or more satellites, and then returns to the same CDC. Multimodality is represented by different transport modes for the first tier, including different types of urban vehicles. As Fontaine et al. (2021), we differentiate between line-based modes (e.g., tram) that can only drive along predefined lines and free-roaming modes (e.g., truck). Satellites represent the second-tier facilities and are usually located near the city center. Here, the load is again consolidated before the final last-mile delivery to the demand location takes place, using city freighters. These city freighters start at a satellite, visit several demand locations and return back to the satellite. A major challenge in these systems is the synchronization between the first-tier operations and the second-tier operations.

3.2 Tactical planning in a 2T-CLS

According to Crainic (2000) and Crainic and Kim (2007), tactical planning aims to develop a transportation plan to facilitate efficient operations and resource utilization while satisfying the demand for transportation within the quality criteria publicized or agreed upon with the respective customers. The tactical plan is designed for a short time horizon, called schedule length, which corresponds to the length of regularity for parameter setting. It must be set up sometime before based on forecasted parameters. Further, we assume a known deterministic demand, each demand associated with a specific volume. Daily demand fluctuations and uncertainties are carried out on the operational level and are therefore out of the scope of our paper.

The tactical plan of a 2T-CLS determines which services to operate, how to allocate demands to those services and satellites, and addresses second-tier routing. The goal is to satisfy the regular demand most efficiently in terms of a cost-efficient and environmentally friendly use of resources, and at the same time to meet demand conditions such as release and due dates. Strategic decisions, such as the location of satellites, are assumed to be given and not addressed in tactical planning. A key challenge within these two-tier systems is to ensure that the tactical plan enables the synchronization of operations between the first and second tiers at the operational level. This synchronization involves precise coordination in terms of location and timing. Concretely, the satellite to which the urban vehicle delivers a demand determines the starting position for that demand on the second tier. Further, in terms of timing, the delivery to a satellite must precede the departure from that satellite of each demand.

3.3 Cooperation in a 2T-CLS

We assume that multiple LSPs operate within the city using the same CDCs and satellites. These LSPs form a coalition and cooperate on a voluntary basis. Each LSP has its own set of demands, each with a certain demand volume, which is brought into the coalition when they join. On the other side, each LSP has its own resources for transportation. These resources are first-tier services, capacities at satellites that they are allowed to use, as well as vehicle fleets. We assume that these LSPs agree to cooperate with each other through a central coordinator responsible for deciding which resources to use in order to fulfill all demands. The central coordinator can either be a third-party provider or a coordinating entity that is jointly operated by the LSPs involved. In either case, each LSP must be willing to share its information about demands and transportation resources.

We assume that the participating LSPs can offer transportation capacities on both the first tier and second tier. Thus, each LSP could satisfy its own demand. Therefore,

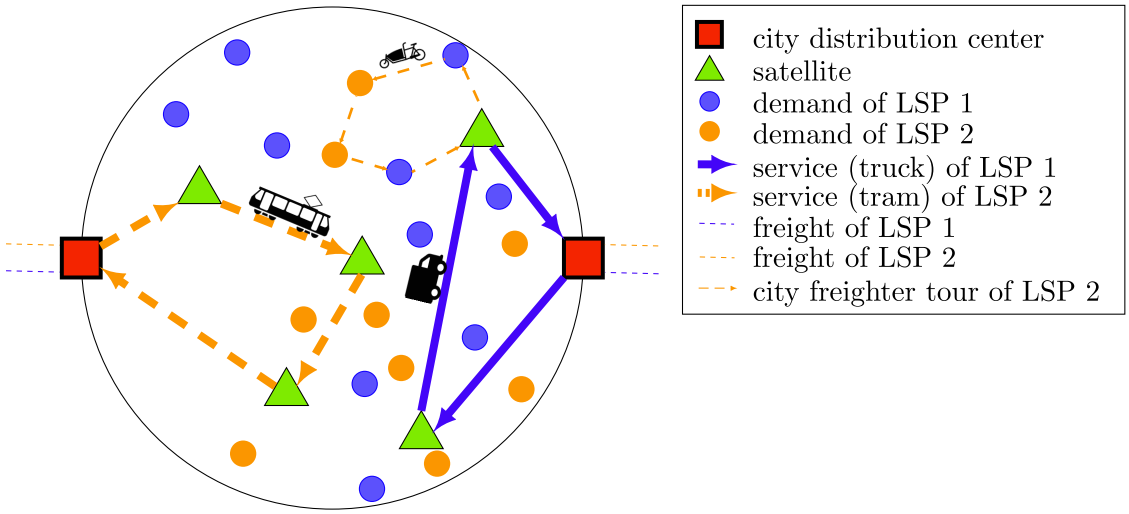

Figure 1: 2T-CLS with cooperating LSPs

cooperation can take place in two possible ways. First, by sharing their customer demands, so that each LSP can also have their demands delivered by services of other LSPs and vice versa. Second, by utilizing the same resources such as vehicle fleets, satellite depots, and CDCs. This enables freight consolidation of the participating LSPs by using common transport resources for first- and second-tier delivery. Figure 1 represents such a 2T-CLSs in which two LSPs cooperate with each other.

Within cooperations, it is important not only that the total costs resulting from the tactical planning of the provided resources are low but also that each individual LSP benefits from the cooperation and perceives the distribution of costs as fair. Therefore, a cost allocation method is required that takes the results of tactical planning as input and allocates the total system costs to the LSPs involved in order to enable fair and stable cooperation.

4 General approach and model

To model the problem at hand, we present a formulation that combines service network design constraints for the first tier and vehicle routing constraints for the second tier. Section 4.1 introduces the general modeling outline and notation. Building upon this, Section 4.2 presents our mathematical formulation for resource planning in a cooperative 2T-CLS.

4.1 General outline and notation

At the core of the problem is a set of LSPs $\mathcal{N}$. Each LSP $n \in \mathcal{N}$ has a subset of customer demands $\mathcal{D}(n)$ out of the set of all demands $\mathcal{D}$. Each demand $d \in \mathcal{D}$ is characterized by a volume v_d, the origin, the destination, the release date that specifies when the demand is available at the CDCs and a due date b_d that specifies when the demand must be delivered to its final location.

Consistent with Crainic et al. (2004), we model the first tier as a service network design. The time dimension is represented by a schedule length that is divided in the periods $1, ..., |\mathcal{P}|$. All considered time-related parameters are assumed to be integer multiples of the period length. Let $\mathcal{E}$ be the set of CDCs and $\mathcal{S}$ the set of satellites. For the first-tier deliveries, a set of services $\mathcal{R}$ is available, characterized by transportation mode m_r, vehicle type t_r and cost c_r, representing not only monetary but also environmental impact cost. Each service starts at a CDC $e_r \in \mathcal{E}$ and visits an ordered sequence of satellites. Transportation modes are represented by the set $\mathcal{M}$. Each mode of transportation is associated with different vehicle types (e.g., the mode "truck" can consist of the types "small truck" and "big truck"), where $\mathcal{T}$ represents the set of vehicle types. Each type t has a specific capacity u_t and a specific fleet size h_{etn} at CDC e provided by LSP n.

We consider three different capacity constraints on satellites. Operations at satellite $s \in \mathcal{S}$ are limited by a maximum number of urban vehicles a_{spn} LSP n can transfer in period p. The same type of constraint is additionally considered by $\bar{a}_{spmn}$ for each mode m to account for different capacity constraints per mode. In addition, we consider an upper limit for the total volume of freight g_{spn} LSP n can accommodate at satellite s in period p. The costs incurred to deliver demand d from its origin to a CDC e are also externally given and denoted as f_{de}. Note that these costs could further include reallocation costs of demands to CDCs of other LSPs if CDCs are not operated by all LSPs together.

To account for the different LSPs, each LSP n is able to provide a subset of services $\mathcal{R}(n) \subseteq \mathcal{R}$. Thereby $\mathcal{R}(n) \cap \mathcal{R}(\hat{n}) = \emptyset \;\; \forall n, \hat{n} \in \mathcal{N} : \hat{n} \neq n$ holds. Further, each service has a certain service time w_r for unloading operations at each satellite, as well as an arrival time τ_{rs} at satellite s, respectively. $\mathcal{R}(p, s)$ and $\mathcal{R}(p, s, m)$ define the subsets of $\mathcal{R}$ that include all services (of mode m) that operate at satellite s during period p, while $\mathcal{R}(p, e, t)$ represents the subset that includes services of vehicle type t, starting at CDC e and operating in period p. We consider the subset $\mathcal{R}(d, s)$ that represents the services that fulfill the release date for demand d at the CDC the service starts from, and further include satellite s in their route.

We include the second-tier routing in our model rather than relying on approximations. To achieve this, we integrate a Vehicle Routing Problem with Release and Due Dates (VRPRDD) (Shelbourne et al., 2017) on each satellite, where the release dates of each demand are determined by the arrival period of the service that the demand is assigned to at the corresponding satellite resulting from the first-tier, plus a service time w_r. For the second-tier routing, we construct a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{A})$ where $\mathcal{V} = \mathcal{D} \cup \mathcal{S} \cup \hat{\mathcal{S}}$ are the nodes consisting of the set of demand locations $\mathcal{D}$, satellite nodes $\mathcal{S}$ that serve as the starting points of the city freighters and a duplication of the satellite nodes $\hat{\mathcal{S}}$ that serve as the ending points of the city freighters. $\mathcal{K}$ represents the set of city freighters. Each city freighter k has a satellite node $s_k^+ \in \mathcal{S}$ at which it starts and a satellite node $s_k^- \in \hat{\mathcal{S}}$ at which it ends. These two satellite nodes are identical, i.e., they refer to the same physical location, as each city freighter starts and ends at the same satellite. The subset $\mathcal{K}(s)$ defines the city freighters starting from satellite s. The set $\mathcal{A}$ represents the arcs connecting the nodes. The subsets $\mathcal{A}(k)$ represent the set of arcs that city freighter k can travel, as each city freighter is only allowed to start and end at the satellite it is assigned to. We specify the allowable arcs for city freighter k to end in node i as $\delta_k^-(i)$, and the arcs starting from node i as $\delta_k^+(i)$. The costs that arise if arc $(i,j) \in \mathcal{A}$ is used by a city freighter are represented by $\hat{c}_{ij}$ and the time distance of each arc $(i,j) \in \mathcal{A}$ is represented by $\bar{t}_{ij}$. We assume service times l of city freighters at each node. Each city freighter has a capacity of q. Again, we account for different LSPs through subsets of city freighter $\mathcal{K}(n)$ provided by LSP n.

Additionally, we define two parameters, α_1 and α_2. These parameters specify for each LSP the lower bound of demand volume that must be assigned to own services on the first tier (α_1) and to own city freighters on the second tier (α_2), respectively. For example, $\alpha_1 = 0.2$ and $\alpha_2 = 0.3$ imply that at least 20% of an LSP's demand volume must be assigned to its own services and at least 30% of an LSP's demand volume must be fulfilled by its own city freighters. Therefore, these two parameters limit the demand sharing between the LSPs. A full table with the notation is provided in A.

4.2 Mathematical problem formulation

In this section, we introduce the MILP for our problem setting in which a central coordinator minimizes the total system costs. After presenting the decision variables and the objective function, we start with the first-tier service network design constraints. Then, we describe the second-tier vehicle routing constraints. On both tiers, we introduce constraints that limit the sharing of demands according to the parameters α_1 and α_2. Further, we introduce linking constraints, which connect the first-tier with the second-tier and enable synchronization on the operational level.

We consider the following decision variables: the binary variable y_r takes the value one if service $r \in \mathcal{R}$ is selected, the binary variable x_{dsr} takes the value one if demand $d \in \mathcal{D}$ is assigned to satellite $s \in \mathcal{S}$ and service $r \in \mathcal{R}$ and the binary variable z_{ijk} takes the value one if second-tier city freighter k uses arc $(i,j) \in \mathcal{A}(k)$. Finally, p_{ik} is specifying

the arrival time of city freighter k at vertex $i \in \mathcal{D} \cup \{s_k^+, s_k^-\}$.
Grounded on this, we formulate the following MILP:

Objective function

$$min \qquad \sum_{r \in \mathcal{R}} c_r \cdot y_r + \sum_{d \in \mathcal{D}} \sum_{s \in \mathcal{S}} \sum_{r \in \mathcal{R}} f_{de_r} \cdot x_{dsr} + \sum_{k \in \mathcal{K}} \sum_{(i,j) \in \mathcal{A}(k)} \hat{c}_{ij} \cdot z_{ijk} \tag{1}$$

The objective function (1) minimizes the total cost incurred. These costs are made up of three components: 1. the costs related to the operation of services, 2. the costs for assigning demands to services and thus simultaneously to the CDCs from which the services start, and 3. the second-tier costs associated with routing the demands from the satellites to their final location.

First-tier service network design constraints

$$\sum_{s \in \mathcal{S}} \sum_{r \in \mathcal{R}} x_{dsr} = 1 \qquad \forall d \in \mathcal{D} \tag{2}$$

$$\sum_{s \in \mathcal{S}} \sum_{r \in \mathcal{R} \setminus \mathcal{R}(d,s)} x_{dsr} = 0 \qquad \forall d \in \mathcal{D} \tag{3}$$

$$\sum_{d \in \mathcal{D}} \sum_{s \in \mathcal{S}} v_d \cdot x_{dsr} \leq u_{t_r} \cdot y_r \qquad \forall r \in \mathcal{R} \tag{4}$$

$$\sum_{r \in \mathcal{R}(p,e,t)} y_r \leq \sum_{n \in \mathcal{N}} h_{etn} \qquad \forall e \in \mathcal{E}, t \in \mathcal{T}, p \in \mathcal{P} \tag{5}$$

$$\sum_{d \in \mathcal{D}} \sum_{r \in \mathcal{R}(p,s)} v_d \cdot x_{dsr} \leq \sum_{n \in \mathcal{N}} g_{spn} \qquad \forall s \in \mathcal{S}, p \in \mathcal{P} \tag{6}$$

$$\sum_{r \in \mathcal{R}(p,s)} y_r \leq \sum_{n \in \mathcal{N}} a_{spn} \qquad \forall s \in \mathcal{S}, p \in \mathcal{P} \tag{7}$$

$$\sum_{r \in \mathcal{R}(p,s,m)} y_r \leq \sum_{n \in \mathcal{N}} \bar{a}_{spmn} \qquad \forall s \in \mathcal{S}, p \in \mathcal{P}, m \in \mathcal{M} \tag{8}$$

$$\alpha_1 \cdot \sum_{d \in \mathcal{D}(n)} v_d \leq \sum_{d \in \mathcal{D}(n)} \sum_{s \in \mathcal{S}} \sum_{r \in \mathcal{R}(n)} x_{dsr} \cdot v_d \qquad \forall n \in \mathcal{N} \tag{9}$$

$$x_{dsr} \in \{0,1\} \qquad \forall d \in \mathcal{D}, s \in \mathcal{S}, r \in \mathcal{R} \tag{10}$$

$$y_r \in \{0,1\} \qquad \forall r \in \mathcal{R} \tag{11}$$

Constraints (2) and (3) guarantee that each demand is assigned to a single service for which the release date at the CDC is not violated. The capacity limit for each urban vehicle type of each service is ensured by Constraints (4). Constraints (5) limit the maximum number of used urban vehicle types t at CDC e for each period p. Constraints (6) limit the amount of freight operated in period p at satellite s. The number of urban vehicles operating in period p at satellite s is limited by Constraints (7) while Constraints (8) explicitly limit the number of urban vehicles of mode m. To limit the demand sharing, Constraints (9) enforce that each LSP satisfies at least a fraction α_1 of its own demand volume by its own services.

Second-tier vehicle routing constraints

For the resulting VRPRDD on the second tier, we use the previously defined graph $\mathcal{G}$.

$$\sum_{k \in \mathcal{K}} \sum_{j \in \delta_k^+(d)} z_{djk} = 1 \qquad \forall d \in \mathcal{D} \qquad (12)$$

$$\sum_{j \in \delta_k^+(s_k^+)} z_{s_k^+ jk} = 1 \qquad \forall k \in \mathcal{K} \qquad (13)$$

$$\sum_{j \in \delta_k^-(d)} z_{jdk} - \sum_{j \in \delta_k^+(d)} z_{djk} = 0 \qquad \forall d \in \mathcal{D}, k \in \mathcal{K} \qquad (14)$$

$$\sum_{i \in \delta_k^-(s_k^-)} z_{is_k^- k} = 1 \qquad \forall k \in \mathcal{K} \qquad (15)$$

$$\sum_{d \in \mathcal{D}} \sum_{j \in \delta_k^+(d)} v_d \cdot z_{djk} \leq q \qquad \forall k \in \mathcal{K} \qquad (16)$$

$$p_{ik} + \bar{t}_{ij} + l - p_{jk} \leq M_1 \cdot (1 - z_{ijk}) \qquad \forall (i,j) \in \mathcal{A}(k), k \in \mathcal{K} \qquad (17)$$

$$p_{dk} \leq b_d \qquad \forall k \in \mathcal{K}, d \in \mathcal{D} \qquad (18)$$

$$\alpha_2 \cdot \sum_{d \in \mathcal{D}(n)} v_d \leq \sum_{d \in \mathcal{D}(n)} \sum_{k \in \mathcal{K}(n)} \sum_{j \in \delta_k^+(d)} z_{djk} \cdot v_d \qquad \forall n \in \mathcal{N} \qquad (19)$$

$$z_{ijk} \in \{0, 1\} \qquad (i,j) \in \mathcal{A}(k), \forall k \in \mathcal{K} \qquad (20)$$

Constraints (12) ensure that each demand location is visited by exactly one city freighter. Constraints (13) to (15) guarantee that each city freighter starts and ends at the respective satellite the city freighter belongs to while ensuring the correct flow. Constraints (16) limit the load of each city freighter to the maximum capacity. Constraints (17) ensure the correct time-flow of each city freighter route, while Constraints (18) guarantee that each demand location is delivered before its due date. Similar to Constraints (9), Constraints (19) limit the demand sharing on the second tier.

Connection constraints

We connect the two tiers with the following constraints:

$$\sum_{r \in \mathcal{R}} x_{dsr} = \sum_{k \in \mathcal{K}(s)} \sum_{j \in \delta_k^+(d)} z_{djk} \qquad \forall d \in \mathcal{D}, s \in \mathcal{S} \tag{21}$$

$$\sum_{r \in \mathcal{R}} \sum_{s \in \mathcal{S}} \tau_{rs} \cdot x_{dsr} + w_r \leq p_{s_k^+ k} + M_2 \cdot (1 - \sum_{j \in \delta_k^+(d)} z_{djk}) \quad \forall d \in \mathcal{D}, k \in \mathcal{K} \tag{22}$$

Constraints (21) state that each demand must be assigned to a city freighter that departs from the same satellite the demand is released to by the chosen first-tier service. Constraints (22) ensure that for each demand, the second-tier departure period at the satellite must be greater or equal to the first-tier arrival period plus service time. The incorporation of these constraints forms the basis for the synchronization of the two tiers at the operational level.

5 Solution approach

As our problem formulation combines two NP-hard problems, we propose a metaheuristic to address large instances. Therefore, we develop an Integrative Two-Step Large Neighborhood Search (I2S-LNS) with adaptive components that integrates decisions on both tiers. We first describe the solution representation and search space in Section 5.1. Afterward, we illustrate the general framework of the metaheuristic in Section 5.2. In Section 5.3, we show how we construct an initial feasible solution. Sections 5.4 and 5.5 describe the two steps of the procedure.

5.1 Solution representation and search space

A solution $\mathcal{S}$ is represented by a set of selected services $\mathcal{R}_s$ (referred to as the service design), the assignment of demands to services and satellites $\mathcal{X}$, and the second-tier routing $\mathcal{Y}$ capturing the sequence of nodes including information on starting and departure times at each visited node. During the entire procedure, the search space is limited to the space of feasible solutions.

5.2 General outline of the I2S-LNS

The general procedure of the I2S-LNS is depicted in Algorithm 1. We start by constructing an initial feasible solution (Section 5.3). Then, our heuristic consists of iteratively applying a two-step procedure, a Large Neighborhood Search on the Service Design (SD-LNS) in Step 1 and an Adaptive Large Neighborhood Search on the Demand Assignment and Routing (DAR-ALNS) in Step 2. Further, we introduce a solution memory to intensify the search in promising search regions.

Precisely in Step 1, we destroy and repair our solution by applying operators on the service design $\mathcal{R}_s$ and adapting $\mathcal{X}$ and $\mathcal{Y}$ to ensure feasibility. Since it is unlikely to find new best solutions with a service design that has high service operating costs, denoted as $c^r(\mathcal{R}_s)$, we employ a threshold $(1+\theta_1) \cdot c^r(\mathcal{R}^*)$ where $c^r(\mathcal{R}^*)$ represents the cheapest service operating cost that are found so far during the search process. If the service operating cost of the current solution exceeds this threshold, we repeat Step 1. If this happens ϕ times in a row, we randomly select one of the stored solutions in our solution memory Γ and go back to this solution. This solution memory is constantly updated during the search process and includes the $|\Gamma|$ best solutions with distinct service design $\mathcal{R}_s$. It plays a crucial role in relocating the search process back to the regions that have proven to lead to high-quality solutions and thus helps to intensify the search in those regions. In Step 2, we destroy and repair our solution by applying operators on the demands. Consequently, the demand assignment $\mathcal{X}$ and routing $\mathcal{Y}$ changes while keeping $\mathcal{R}_s$ fixed. In both steps, the best solution $\mathcal{S}^*$ is always updated when a new best solution is found. We stop after a minimum number of iterations without improving the objective value of the best found solution $f(\mathcal{S}^*)$ or after a time limit is reached.

Algorithm 1 I2S-LNS

```
S_c ← Generate a feasible initial solution;                    Section 5.3
while Stopping criterion not met do
    it ← 0;
    repeat
        if it > φ then
            S_c ← drawRandomElement(Γ);                        Use solution memory
            break
        end
        S_c ← SD-LNS(S_c);                                     Step 1: Section 5.4
        it ← it + 1;

    until c^r(R_s) < (1 + θ_1) · c^r(R*);
    S_c ← DAR-ALNS(S_c);                                       Step 2: Section 5.5
end
```

Compared to standard large neighborhood searches from the literature, we have both a large neighborhood search for service design and an adaptive large neighborhood search for demand assignment and routing in one procedure. These two main components alternate and are improved in their coordination and efficiency by the intelligent inclusion of thresholds and a solution memory. This allows us to utilize the existing two-tiered problem structure efficiently.

5.3 Construction heuristic

We sequentially generate a feasible initial solution for the first tier and then for the second tier and ensure connection

First tier: On the first tier, the solution consists of $\mathcal{R}_s$ and $\mathcal{X}$. In a first step, a service that complies with all the capacity restrictions (see Constraints (5) to (8)) is randomly selected. Afterward, demands that are not assigned to a service yet are randomly drawn one after the other and assigned to the selected service and assigned to the closest possible satellite the service visits, if constraints allow. We repeat this procedure until all demands are assigned. Assignments to satellites and services are also only permitted if it is ensured that there is still enough time for at least a commuting tour of a city freighter from the satellite to the final demand location to ensure feasibility on the second tier. This procedure is repeated η_1 times to generate η_1 different starting solutions. The solution with the lowest costs is taken as the initial solution for the first tier.

Second tier: Resulting from the first-tier assignments to satellites, we have a VR-PRDD on each satellite. For each satellite, we greedily insert each demand in its cheapest possible position until all demands are served. This generates the initial solution for the routing $\mathcal{Y}$.

5.4 Step 1: SD-LNS

We start by destroying the current solution S_c by applying removal operators on the service design $\mathcal{R}_s$. As the solution gets infeasible when removing services because previously assigned demands are getting unassigned, we try to reassign the unassigned demands to other services in the existing service design in their cheapest possible position (in terms of demand assignment to services and satellites as well as routing). If not all demands can be reassigned into the existing service design, we insert new services using the insertion operators. We repeat this until all demands are assigned to a service, resulting in a fully repaired solution. Algorithm 2 depicts the complete logic of the SD-LNS.

For removing services, we use the following operators:

- **Random removal**: We randomly select between 1 and $\lambda_{\mathcal{R}}$ services and remove them.

Algorithm 2 SD-LNS

$\mathcal{S}_c \leftarrow removalOperator(\mathcal{S}_c)$
$\mathcal{D}_u \leftarrow$ Demands that have been assigned to removed services
for $d \in \mathcal{D}_u$ **do**
  if $bestInsertion(d, \mathcal{S}_c)$ **then**
    $\mathcal{D}_u \leftarrow \mathcal{D}_u \setminus \{d\}$;
  end
end
while $D_u \neq \emptyset$ **do**
  $\mathcal{S}_c \leftarrow insertionOperator(\mathcal{S}_c)$;
  for $d \in \mathcal{D}_u$ **do**
    if $bestInsertion(d, \mathcal{S}_c)$ **then**
      $\mathcal{D}_u \leftarrow \mathcal{D}_u \setminus \{d\}$;
    end
  end
end
$\mathcal{S}^* \leftarrow \text{update}(\mathcal{S}^*, \mathcal{S}_c)$; $\Gamma \leftarrow \text{update}(\Gamma, \mathcal{S}_c)$; $\mathcal{R}^* \leftarrow \text{update}(\mathcal{R}^*, \text{S}_c)$;

- **Worst removal**: For each currently active service, we determine the ratio between the demand volume that is currently assigned to the service and the service operating cost. Between 1 and $\lambda_{\mathcal{R}}$ services with the worst ratio are selected and closed.

To insert new services, we first determine all services that can still be inserted without violating any capacity constraints. Further, we determine if the service must be provided by a specific LSP to not violate the demand sharing Constraints (9). Then, we apply the following operators to this set of services.

- **Random insertion**: We randomly select a service and add it to the set of selected services $\mathcal{R}_s$.
- **Best-fit insertion**: We select the service with the best ratio of demand volume of the unassigned demands for which the service fulfills the release and the due date, and service operating cost c_r. In a preprocessing step before the heuristic starts, we determine for each demand and satellite combination the subset of services, that fulfill the release and due dates. Through this preprocessing step, we can easily evaluate this operator.

For continuous diversification purposes, we multiply the sorting criteria for the worst removal and the best-fit insertion operators by a random number in the interval [0.8, 1.2]. The selection of operators is based on the probabilities μ_r and μ_i, which determine the probability of utilizing the random operator for removal and insertion, respectively. The calibration of these probabilities is detailed in Section 7.2.1.

5.5 Step 2: DAR-ALNS

In this step, we iteratively destroy and repair the solution of Step 1 by applying operators on the demands. Throughout this step, the service design $\mathcal{R}_s$ is fixed. Only the demand assignment $\mathcal{X}$ and routing $\mathcal{Y}$ are modified. In each iteration, we first choose a demand removal operator to destroy our current solution. After that, we repair our solution by applying an insertion operator. We only update the current solution if the objective value is not increased by more than a factor of $(1+\theta_w)$. We again take advantage of a threshold to prevent wasting computation time in unpromising search regions. Specifically, we break the DAR-ALNS if the objective value of the current solution $f(\mathcal{S}_c)$ exceeds a dynamic threshold that starts in the first iteration with $(1+\theta_2)\cdot f(\mathcal{S}^*)$ and decreases linearly with the number of iterations until it reaches $f(\mathcal{S}^*)$. We employ this logic since strong improvements in the current solution can be expected, especially during early iterations. We set a maximum number of iterations η_2 as stopping criteria. Algorithm 3 depicts the complete logic of the DAR-ALNS.

Algorithm 3 DAR-ALNS

$it \leftarrow 0$;
while $it < \eta_2$ **do**
  $\mathcal{S}'_c \leftarrow Repair(Destroy(\mathcal{S}_c))$ through applying operators on $\mathcal{D}$
  $it \leftarrow it + 1$;
  if $f(\mathcal{S}'_c) < (1+\theta_w)\cdot f(\mathcal{S}_c)$ **then**
    $\mathcal{S}_c \leftarrow \mathcal{S}'_c$; update current solution
  end
  $\mathcal{S}^* \leftarrow$ update$(\mathcal{S}^*, \mathcal{S}_c)$; $\Gamma \leftarrow$ update$(\Gamma, \mathcal{S}_c)$;
  if $f(\mathcal{S}_c) > (1+\theta_2\cdot(1-it/\eta_2))\cdot f(\mathcal{S}^*)$ **then**
    break; break if threshold is exceeded
  end
end

We consider the following problem-specific operators to remove demands from the solution:

- **Random removal**: We randomly draw between 1 and $\lambda_{\mathcal{D}}$ demands.
- **CDC regret removal**: For each demand d, we determine whether the assignment to another CDC e would lead to lower costs f_{de}. We then sort the demands according to the potential cost savings and take between 1 and $\lambda_{\mathcal{D}}$ demands with the highest potential cost savings when assigned to another CDC.
- **Satellite regret removal**: For each demand, we determine the deviation between the distance to the satellite the demand is assigned to and the distance to the closest other satellite. We sort the demands increasing according to this deviation and select between 1 and $\lambda_{\mathcal{D}}$ demands with the highest deviation. The idea behind this operator is that demands that are not assigned to their closest satellite are more likely to lead to cost savings.

- **Random route removal**: A city freighter route is randomly selected and all the demands that are on this city freighter are removed.
- **Worst route removal**: For each city freighter route, we determine the ratio between the routing costs and the demand volume that is fulfilled by this city freighter. All demands of the city freighter route with the worst ratio are removed.

For each of the operators that are not random, we multiply the sorting criteria by a random number in the interval [0.8, 1.2] for better continuous diversification.

We consider the following insertion operator

- **Best insert**: We insert the demand where the costs, in terms of service assignment cost as well as second-tier routing insertion cost, are lowest.
- **2-regret insert**: For each demand, we determine the cheapest insertion cost for assigning the demand to each service and satellite combination. The regret value is then defined by the difference between the cheapest and the second cheapest service / satellite combination. Then we sort the demands in decreasing order by their regret value and insert them one after the other in their cheapest possible position in terms of service assignment cost as well as second-tier routing cost.

If, for insertion purposes, a new city freighter must start, we assign the city freighter to the LSP of the respective demand.

Operator selection Operators are chosen based on a roulette wheel mechanism, as described by Pisinger and Ropke (2007), utilizing the reaction factor ρ. Each pairing of removal and insertion operators is assigned a reward σ, allocated as follows: σ_1 for worsening the current objective value, σ_2 for maintaining the current objective value, σ_3 for improving the current objective value. Unlike many studies, we do not assign a reward for finding a new best global solution because such outcomes are significantly influenced by the current service design $\mathcal{R}_s$, which is not impacted by these operators.

6 Cost allocation in a cooperative 2T-CLS

After minimizing the total system costs, the question now arises of how to allocate these costs to the participating LSPs. Within our study, we consider proportional methods (Section 6.1), the Shapley Value (Section 6.2), and the EPM (Section 6.3). We focus on these two game theoretical methods, as they represent two different approaches, one based on marginal costs (Shapley Value) and the other based on similar relative savings (EPM).

6.1 Proportional methods

The methods most commonly used in practice because of their simplicity are proportional methods. Proportional methods distribute costs in proportion to a predefined reference value (Vanovermeire et al., 2014). Although these methods are easy to calculate and interpret, they only take one aspect into account and leave out other key elements that influence the cost of each LSP.

In our work, the main cost drivers are the number of demands as well as the demand volume. Therefore, we include the following problem-specific cost allocation methods.

- **Demand-based Allocation (DA)**: The cost allocation to LSP n is proportional to the number of demands.
$$DA_n = \frac{|\mathcal{D}(n)|}{|\mathcal{D}|} \tag{23}$$
- **Volume-based Allocation (VA):** The cost allocation is proportional to the total volume of demand.
$$VA_n = \frac{\sum\limits_{d \in \mathcal{D}(n)} v_d}{\sum\limits_{d \in \mathcal{D}} v_d} \tag{24}$$
- **Demand- and Volume-based Allocation (DVA):** The DVA is calculated as a weighted sum (weight ω) of the DA and VA where $0 \leq \omega \leq 1$.
$$DVA_n = \omega \cdot DA_n + (1 - \omega) \cdot VA_n \tag{25}$$
- **Stand-Alone-based Allocation (SAA)**: The cost allocation is determined by the ratio of the stand-alone costs of each LSP $C(\{n\})$ when no cooperation occurs and the cumulated stand-alone cost of all LSPs. In contrast to the other proportional methods, however, it is necessary to calculate the costs that incur for each LSP when acting independently.
$$SAA_n = \frac{C(\{n\})}{\sum\limits_{\hat{n} \in \mathcal{N}} C(\{\hat{n}\})} \tag{26}$$

Each of the four cost allocations is interpreted as a share of the total costs that LSP n must bear.

6.2 Shapley Value

Shapley (1953) introduced the Shapley Value as a cost allocation method in cooperative game theory. The Shapley Value is calculated as the average marginal contribution that the participant makes to each possible subcoalition when added to the coalition according to the order. It provides a unique cost allocation solution that satisfies a lot of fairness axioms. For details about the formulation of these axioms, we refer to Shapley (1953). Analytically, the Shapley Value for a participant n is calculated as follows:

$$SV_n = \sum_{\mathcal{O} \subset \mathcal{N} : n \in \mathcal{O}} \frac{(|\mathcal{O}| - 1)!(|\mathcal{N}| - |\mathcal{O}|)!}{|\mathcal{N}|!} \cdot [C(\mathcal{O}) - C(\mathcal{O} - \{n\})] \tag{27}$$

Thereby, the summation is over all subcoalitions $\mathcal{O}$ within the set of all participants $\mathcal{N}$ that contain participant n whereby $[C(\mathcal{O}) - C(\mathcal{O} - \{n\})]$ represents the participants marginal contribution to the subcoalition $\mathcal{O}$. The Shapley Value is therefore a marginal cost-based method that does not emphasize an equal distribution of costs but derives its fairness aspect from marginal costs.

6.3 Equal Profit Method

The EPM was developed by Frisk et al. (2010) and is based on the idea that the maximum difference of the pairwise relative savings should be minimized. Thus, the relative savings of the participants should be as similar as possible. With decision variable c_n representing the costs allocated to participant n and $C(\{n\})$ representing the stand-alone costs of participant n the EPM allocates the costs through the following linear model:

$$\min \quad f \tag{28}$$

$$\text{s.t.} \quad f \geq \frac{c_n}{C(\{n\})} - \frac{c_{\hat{n}}}{C(\{\hat{n}\})} \qquad \forall n, \hat{n} \in \mathcal{N} : n \neq \hat{n} \tag{29}$$

$$\sum_{n \in \mathcal{O}} c_n \leq C(\mathcal{O}) \qquad \forall \mathcal{O} \subset \mathcal{N} \tag{30}$$

$$\sum_{n \in \mathcal{N}} c_n = C(\mathcal{N}) \tag{31}$$

The objective function (28) minimizes the maximum pairwise difference in relative savings that is measured by Constraints (29). Constraints (30) and (31) ensure the rationality and the efficiency condition. The constraints ensure a core allocation if a cost allocation that lies in the core exists. In the case that the core is empty, we use the epsilon-core as proposed by Frisk et al. (2010) to keep the coalition stable.

7 Numerical study

Section 7.1 describes the numerical setup and the generation of instances. Then, we calibrate the parameters and evaluate the performance of the I2S-LNS in Section 7.2. Section 7.3 presents managerial insights regarding the impact of cooperation as well as about the before-introduced cost allocation methods.

7.1 Numerical setup and instance generation

We implemented the MILP in Python 3.12 using Gurobi 11 as a solver. The I2S-LNS is implemented in C++. All experiments are carried out on an AMD Ryzen 9 5950X 16-Core Processor, 3.40 GHz with 128 GB RAM. We generate new specific instances for this problem. Similar to related papers in the field of 2T-CLSs, we generate the instances based on a real city (Fontaine et al., 2021; Crainic et al., 2004). In our case, we take a large German city as a basis. We locate CDCs in easily accessible places outside the city. For satellites, we mainly use tram stops or larger squares in the city center. We go one step further than other papers and sample demand destinations for all instances in relation to publicly available data on population density in different city districts. We use two different networks, hereinafter called N1 and N2. N1 has two CDCs and four satellites, N2 has three CDCs and six satellites. We consider a planning horizon of 36 periods, ten minutes each. We generate a set of services with different numbers of satellites starting and ending at a CDC. The first start period is chosen randomly between periods five and ten. Each drawn service is then replicated two times during the planning horizon, with eight periods in between. We consider four different urban vehicle types: small tram, large tram, small truck, and large truck, with capacities of 500 for small vehicles and 750 for large vehicles. The cost of the services c_r depends on the vehicle type and the travel distance. We set the fixed cost for small vehicles to 15 and for large vehicles to 20. The variable costs are the travel distance multiplied by 1.2 for small trams, 1.5 for small trucks, 1.7 for large trams, and 2.0 for large trucks. The lower cost of trams is justified by their lower environmental impact cost. Further, the travel speed for trams is set to 25 km/h and for trucks to 20 km/h because of traffic. The capacity of city freighters on the second tier is 250, and the travel speed is 20 km/h. A service time of one period is assumed. In each instance, each LSP individually draws random services out of the generated services, with half of these services being operated by large urban vehicles. Demand volumes are uniformly distributed between 50 and 100. For each demand, the release date RD is randomly selected from the range 1 to 18. The due date DD is calculated using the formula $DD = RD + 12 + \text{randint}(0, 6)$ to ensure time feasibility. For each demand / CDC combination, we assume random assignment cost (f_{de}) ranging from one to five. For capacity constraints, we assume that each LSP is allowed to operate one vehicle at a time (a_{spn}) at each satellite, implying $a_{spnm} \leq 1$ for all modes, and use 300 units of the satellite capacity (g_{spn}). With this setting, we generate instances of different sizes with up to three LSPs, 100 demands, and 60 services for our numerical experiments in the following sections.

7.2 I2S-LNS performance

We start with calibrating the parameter setting of the I2S-LNS (Section 7.2.1). No benchmark instances in the literature exist integrating service network design on the first tier and vehicle routing on the second tier. Further, comparing to pure service network design problems or pure two-echelon vehicle routing problems does not consider essential parts of our problem setting. Therefore, we evaluate the performance of the I2S-LNS through three steps: First, by comparing it against Gurobi on small instances (Section 7.2.2), second, by analyzing its performance on larger instances (Section 7.2.3), lastly, by measuring the performance impact of removing newly introduced components of the I2S-LNS (Section 7.2.4). Throughout all performance experiments, we assume two LSPs that are fully allowed to share their customers ($\alpha_1 = \alpha_2 = 0$). Demands and services are evenly split across these two LSPs.

7.2.1 Calibration

To calibrate the parameter setting, we conducted tests on medium-sized instances with 40 and 60 demands. Initial values for each parameter were determined based on preliminary tests conducted throughout the development phase, ensuring a realistic starting point. Subsequently, we systematically varied each parameter within a specified range while holding all others constant to identify the setting that yielded the best average performance. The finalized parameter settings that are used throughout our experiments are presented in Table 1.

Table 1: I2S-LNS parameter setting

Parameter	**Value**	**Description**
η_1	100	Number of iterations of first-tier construction heuristic
η_2	$1.1 \cdot \lvert\mathcal{D}\rvert$	Number of iterations in Step 2
θ_1	0.25	Parameter for threshold in Step 1
θ_2	0.35	Parameter for threshold in Step 2
θ_W	0.01	Parameter for threshold for worsening the objective value
λ_D	6	Max. number of removed demands
$\lambda_{\mathcal{R}}$	3	Max. number of removed services
μ_r	0.2	Probability for random removal operator
μ_i	0.4	Probability for random insertion operator
ϕ	5	Max. number of threshold violations in Step 1
$\lvert\Gamma\rvert$	5	Length of solution memory
ρ	0.1	ALNS reaction factor
σ_1, σ_2, σ_3	{1,3,5}	ALNS scores for updating weights

7.2.2 Benchmark against Gurobi

We compare the performance of the I2S-LNS with the solutions obtained by solving the MILP using Gurobi. A time limit of one hour is imposed on Gurobi, whereas the I2S-LNS is set to terminate after 1.000 iterations (full iterations of Step 1 and Step 2) without improvement or once a time limit of ten minutes is reached. We execute the I2S-LNS five times for each instance. We conduct two experiments for this purpose. In the first experiment, we increase the number of demands ($|\mathcal{D}|$) while keeping the number of services constant ($|\mathcal{R}| = 24$). In the second experiment, the number of services ($|\mathcal{R}|$) is increased while the number of demands ($|\mathcal{D}| = 15$) is kept constant. To exclude the influence of the demands on the complexity, corresponding instances share identical demand sets. For example, the first instance with $|\mathcal{R}| = 12$ has the same demands as the first instance with $|\mathcal{R}| = 36$, and this pattern is maintained across all instances. For the first experiment, we use both networks. For the second experiment, we only use N2 to generate a wider range of services. The aggregated results for these experiments are shown in Table 2 and Table 3 (detailed results for all benchmarks are presented in B). We report the best and the average objective values as well as the percentage deviation (σ[%]) between the best and average. Further, we report the percentage difference between the objective value obtained by Gurobi and the average objective value of the I2S-LNS (Δ[%]) and the corresponding runtime.

Table 2: Aggregated comparison of Gurobi and the I2S-LNS with varying $|D|$ and constant $|R| = 24$

		Gurobi			I2S-LNS				
Network	$\|\mathcal{D}\|$	**Costs**	**GAP** [%]	**Time** [s]	**Best costs**	**Avg. costs**	σ [%]	**Avg. time** [s]	Δ [%]
N1	5	108.85	0.0	2	108.85	108.85	0.0	17	0.0
N1	10	169.63	1.17	1731	169.63	169.63	0.0	37	0.0
N1	15	234.33	24.95	3600	226.21	226.55	0.14	107	-3.01
N1	20	285.50	24.22	3600	275.30	275.81	0.19	180	-3.29
N2	30	412.80	33.91	3600	371.58	373.68	0.55	405	-9.32
N2	40	-	-	3600	486.65	490.13	0.65	544	-

In the first experiment, we observe that Gurobi is only able to prove optimality for very small instances with $|\mathcal{D}| = 5$ and $|\mathcal{D}| = 10$. For larger instances, Gurobi exhibits large gaps of up to 33.91% on average for instances with $|\mathcal{D}| = 30$. In instances with $|\mathcal{D}| = 40$, Gurobi fails to obtain a feasible solution within the time limit. Our metaheuristic significantly outperforms Gurobi by finding exactly the same solutions for very small instances where Gurobi finds the optimal solution and better solutions for larger instances in less computation time. Notably, for larger instances, the performance of our metaheuristic remains very stable, with mean σ significantly below 1%.

In the second experiment, we observe a slight increase in the gaps shown by Gurobi as the number of services rises. Notably, the difference between the I2S-LNS and Gurobi

Table 3: Aggregated comparison of Gurobi and the I2S-LNS with varying $|R|$ and constant $|D| = 15$

	Gurobi			I2S-LNS				
$\|R\|$	**Costs**	**GAP [%]**	**Time [s]**	**Best costs**	**Avg. costs**	σ [%]	**Avg. time [s]**	Δ [%]
12	230.01	15.21	3600	226.62	226.62	0.00	93	-1.49
36	222.71	20.56	3600	215.45	215.46	0.00	104	-2.88
60	205.19	24.58	3600	198.87	198.91	0.02	109	-2.98

also increases slightly as the number of services increases. In all instances, the I2S-LNS delivers equal or better solutions than Gurobi in a much shorter computation time. Our metaheuristic also shows remarkable stability, with σ equals 0 in most instances, reflecting minimal variation between the best and average results. The computation time is only slightly influenced by the increased number of services. This analysis highlights the number of demands as the key complexity driver in our problem setting.

7.2.3 Performance on larger instances

In Table 4, we assess the heuristics stability on larger instances, creating five instances each for three configurations: $\mathcal{D} = 50, \mathcal{R} = 36$; $\mathcal{D} = 75, \mathcal{R} = 48$; and $\mathcal{D} = 100, \mathcal{R} = 60$. We set the time limit to 30 minutes.

Table 4: Aggregated performance of the I2S-LNS on larger instances

$\|D\|$	$\|R\|$	**Best costs**	**Avg. costs**	σ [%]	**Avg. time [s]**
50	36	587.16	590.48	0.57	1268
75	48	877.74	884.02	0.71	1778
100	60	1105.47	1113.75	0.75	1800

We observe that even for larger instances, our heuristic delivers very stable results with σ remaining below 1% in almost all of the instances.

7.2.4 Impact of special components

In this section, we assess the performance impact of special components of the I2S-LNS, specifically the thresholds and the solution memory. Table 5 presents the average percentage increase in the objective value when removing the special components. Note that when leaving out threshold θ_1, we also do not take advantage of the solution memory Γ. We use the previously introduced larger instances with 50 to 100 demands.

As can be clearly seen, the inclusion of threshold θ_1 and, thus, the inclusion of the solution memory Γ dramatically boosts the performance of our heuristic among all in-

Table 5: Increase in objective value when taking out special components

$\lvert\mathcal{D}\rvert$	**Threshold θ_1 and memory Γ**	**Threshold θ_2**
50	2.66%	0.25%
75	2.65%	0.79%
100	3.17%	0.75%

stance sizes. Also, the inclusion of the threshold θ_2 significantly contributes to a better performance of our heuristic. These results show that the thresholds in this two-step procedure are enhancing the performance of the I2S-LNS.

7.3 Managerial insights

Firstly, we analyze the impact of varying the cooperation intensity through modifying the $\alpha-$parameters. Subsequently, we compare and analyze the cost allocation methods that were previously introduced. For these investigations, we generate ten instances each with 48 demands distributed across three LSPs, each offering 21 services. We take three cases into account, with each case being a modification of the previous.

- In **Case 1**, demands are evenly split among the LSPs, meaning each LSP has 16 demands.
- In **Case 2**, we adjust Case 1 by changing the allocation of demands to LSPs; one LSP receives eight demands, another 16, and the last one 24 demands.
- In **Case 3**, we further modify Case 2 by also varying the demand volume. The demand volume for the LSP with 16 demands is multiplied by 0.75, and for the LSP with eight demands by 0.5, to assess the effects of altering both the number and the volume of demands.

Thus, we have in Case 1 equally sized LSPs, in Case 2 LSPs that differ by the number of demands, and in Case 3 LSPs that differ by the number of demands as well as by the demand volume. For all experiments LSP 1 refers to the smallest, LSP 2 to the medium and LSP 3 to the largest LSP.

7.3.1 Impact of varying the cooperation intensity

In this section, we examine the effects of the cooperation intensity constraints. Starting from a fully cooperative system in which both resources and demands are fully shared ($\alpha_1 = \alpha_2 = 0$), we gradually increase the α parameters to 1 (no demand sharing). First, we increase α_1 and α_2 both at the same time (Figure 2). Then, we increase them individually while holding the other α parameter constant at 0 (Figure 3).

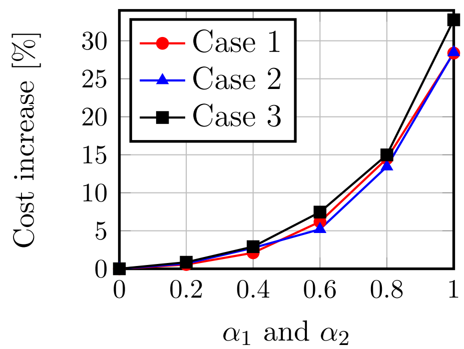

Figure 2: Sensitivity analysis for increasing α_1 and α_2

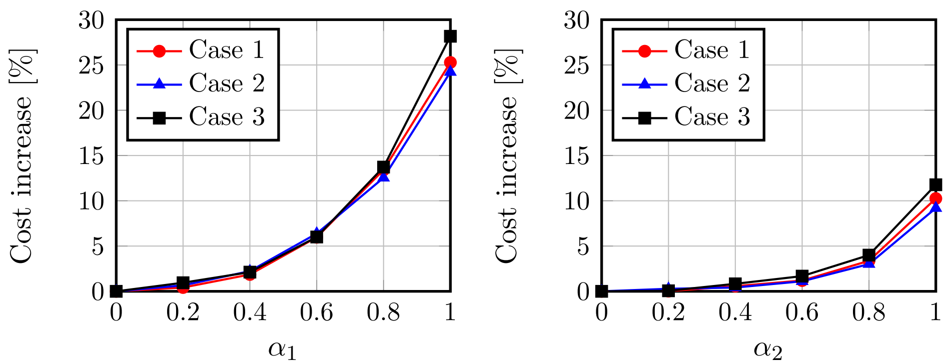

Figure 3: Sensitivity analysis for increasing α_1 and α_2 individualy

We observe that increasing the $\alpha-$parameters to 1 leads to significant increases in total system costs across all cases. On average, total cost increases by 29.25% when there is no demand sharing compared to total demand sharing. Additionally, we find that sharing even 20% of demands ($\alpha = 0.8$) can lead to a large portion of the potential cost savings achievable through full demand sharing. This is especially true for cooperation on the second tier, as nearly the whole increase in cost is happening for $\alpha_2 > 0.8$. This is attributed to the fact that LSPs use identical city freighters, eliminating structural cost differences among these LSPs. Further, we identify that sharing demands in the first tier leads to a higher cost impact than sharing demands in the second tier. This is due to the fact that first-tier cooperation impacts the assignment to CDCs, the utilization of the services, as well as the assignment to the satellites and thus the starting point for the second-tier routing. Moreover, the larger capacity of first-tier urban vehicles enhances the demand pooling potential across different LSPs. These effects result in a very interesting insight for LSPs, as a very large part of the potential cost savings is possible through cooperation without many of the customers even realizing this, as the final delivery to their homes is carried out by the LSPs they have engaged. This is demonstrated by the fact that almost the full amount of potential cost savings can be realized through full cooperation on the first tier and only the sharing of a small part of the demands on the second tier.

Additionally to the cost increase compared to full cooperation, we further investigate the effects of demand sharing on some major KPIs, namely the utilization of first-tier urban vehicles, the utilization of city freighters, and the share of provided capacity by selected services that are operated by large vehicles (large trams and large trucks). The results are presented in Table 6.

Table 6: KPIs for different cooperation scenarios

KPI	**Full coop.**	**First tier coop. only**	**Second tier coop. only**	**No coop.**
Util. first tier vehicles [%]	89.07	89.05	76.13	77.90
Util. second tier vehicles [%]	82.42	69.39	79.07	72.77
Share of large vehicles [%]	60.14	61.14	47.09	43.61
Cost increase [%]	-	10.40	25.89	29.87

Regarding the KPIs, we identify the following effects: 1. Cooperation, in general, increases vehicle utilization. 2. Cooperation leads to a higher share of larger urban vehicles. 3. Cooperation exclusively on the first tier enhances vehicle utilization on this tier while slightly decreasing it on the second tier. This effect is due to pooling demands of different LSPs into a single service. Consequently, diverse demands from these LSPs reach the satellites leading to reduced utilization on the second tier as each LSP is required to operate its own city freighters separately.

These results clearly underline that demand sharing not only leads to monetary savings for the LSPs but also to significantly lower negative impact on the environment.

7.3.2 Cost allocation

In this section we examine the cost savings for the various LSPs when using the introduced cost allocation methods. In particular, the aim is to find out whether there are different interests in the selection of the cost allocation method between the different sized LSPs. For this analysis, we evaluate the characteristic function of each instance by running the heuristic for each possible sub-coalition of LSPs. We then apply the previously introduced cost allocation methods to the individual instances. We use $\omega = 0.5$ as the weight for the DVA.

Our analysis reveals that by fully sharing resources and demands, the coalition could achieve cost savings of on average 26.91% by comparing the total stand-alone costs of all LSPs to the costs of the entire coalition across all cases. To compare the different cost allocation methods, Table 7 shows the average percentage cost savings of each LSP

Table 7: Average cost savings [%] compared to stand-alone cost

Case	LSP	SV	EPM	DA	VA	DVA	SAA
Case 1	1	26.82	26.32	24.59	25.87	25.23	26.32
	2	25.71	26.32	26.96	26.44	26.70	26.32
	3	26.54	26.32	26.98	26.34	26.66	26.32
Case 2	1	41.29	26.36	36.69	38.62	37.65	26.36
	2	25.24	26.36	24.36	24.07	24.21	26.36
	3	21.11	26.36	23.25	25.51	24.38	26.36
Case 3	1	47.31	28.08	33.59	61.77	47.68	28.08
	2	27.84	28.08	21.78	29.77	25.78	28.08
	3	21.41	28.08	29.51	14.82	22.17	28.08

compared to its stand-alone cost using the cost allocation methods. Thereby, LSP 1 stands for the smallest, LSP 2 for the medium, and LSP 3 for the largest LSP.

Based on the results, we identify the following effects for the three cases:

- In Case 1, the average cost savings do not differ significantly neither between the LSPs nor between the individual cost allocation methods. This is not surprising as the LSPs do not differ significantly as they have exactly the same number of demands and services drawn from the same population. Differences at the level of the individual instances balance each other out on average.
- In Case 2, we find that LSP 1 is particularly favored by the use of the demand-based proportional methods and the Shapley Value, as the relative cost savings are significantly higher when using this method than when using the other methods. In contrast, it is better for the two larger LSPs to allocate costs using methods that generate similar relative cost savings. This is due to the fact that larger LSPs usually already have higher utilization of services because of their high number of demands. Demand-based proportional cost allocation methods do not take this effect into account, which leads to lower relative cost savings of larger LSPs compared to smaller LSPs when using these allocation methods.
- In Case 3, we observe similar effects. LSP 1 benefits in particular from the use of the Shapley Value as well as from the use of the VA, as this takes into account both the lower number of demands and the lower demand volume. LSP 3 again benefits, in particular, from the EPM and the SAA, but this time also from the DA, as this method does not take into account the higher demand volume compared to the other LSPs.

Notably, we identify that in every instance, the SAA and EPM deliver the exact same results. This is due to the fact, that allocating costs according to the SAA method already satisfies the additional constraints that are included in the EPM method.

Further, the experiments show that there are diverging interests of the LSPs with regard to the selection of a cost allocation method based on their size. Simple proportional methods may seem plausible and easy to calculate but can lead to very different relative cost savings among the LSPs. In particular, it would be difficult to agree on a fair reference value for the proportional methods, as this leads to immense differences in cost allocation. With respect to the two game theoretical allocation methods, it is particularly desirable for large LSPs to use the EPM or the SAA, as they lead to the same relative cost savings among the LSPs. For the smallest LSP, the allocation according to the Shapley Value is advantageous, as it takes into account the comparatively low marginal costs once the smallest LSP joins a coalition. These results show the diverging interests of the individual LSPs in relation to the selection of a fair allocation method.

Detailed analysis further reveals the following effects:

- In Case 3, the VA results in cost allocations that seem to be unfair for the two larger LSPs. Specifically, in three out of ten instances, the costs allocated to the two larger LSPs exceed the costs of the subcoalition of these two LSPs. In Case 2, we observe a similar effect in the allocation according to the DA. In several cases, this leads to an allocation where the two largest LSPs have almost no cost savings compared to forming a joint subcoalition. This suggests that simple proportional allocation methods can potentially lead to cost allocations that do not allow for stable coalitions, as some subcoalitions are more profitable than the coalition of all LSPs.
- Although the average cost allocations in Case 1 are quite similar, significant differences in individual instances exist, especially between the game theoretical methods and the proportional methods (except SAA). This discrepancy arises because the proportional methods do not account for variations in costs, which are influenced by the location of demands, the composition of the services, and the release and due dates for these demands.

Another important insight that we identify is that the smallest LSP (LSP 1) only incurs a very small marginal cost when added to the coalition of LSP 2 and LSP 3. Figure 4 displays both the stand-alone costs (SAC) of LSP 1 when acting independently and the marginal costs (MC) of adding LSP 1 to the coalition throughout the ten instances for Case 2 and 3. As shown, the marginal costs are by far lower compared to the stand-alone costs. Remarkably, in one instance in Case 3, the marginal costs are even negative. This is attributed to the fact that LSP 1 contributes only a few low-volume demands to the coalition but offers as much capacity at the satellites and provides as many services as the other LSPs. These counteracting cost effects result in very low marginal costs. This clearly indicates that incorporating smaller LSPs, which also have transport capacities, can be highly beneficial for the coalition overall.

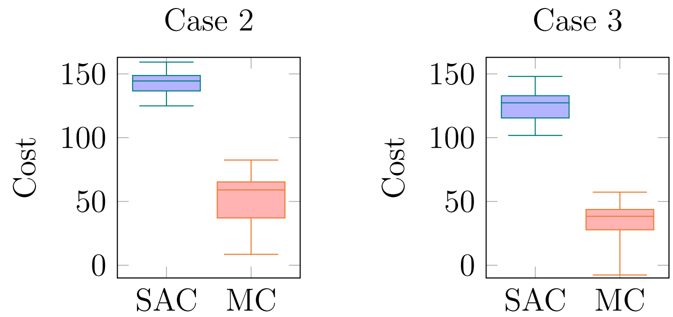

Figure 4: Distribution of stand-alone cost and marginal cost for LSP 1

8 Conclusion and further research

This study presented a decision model for cooperation in 2T-CLSs, which integrates a service network design on the first tier and a vehicle routing problem on the second tier. Additionally, we proposed an efficient problem-specific metaheuristic that is based on a two-step procedure consisting of a large neighborhood search for the service design and an adaptive large neighborhood search for the demand assignment and routing. Furthermore, various cost allocation methods were introduced and applied for this problem setting.

Through an extensive numerical study, we demonstrate the performance of our solution approach. Further, we could show that the cooperation among LSPs can result in significant cost savings. Even when only a small percentage of the demands is shared among the coalition, a large part of the potential cost savings through cooperation can be achieved. Further, cooperation contributes to increased vehicle utilization on both tiers and thereby contributes towards a more sustainable city logistics system. What is particularly important for LSPs is that we were able to show that almost all potential cost savings through full cooperation can be realized through full cooperation on the first tier and the sharing of only a small share of the demands on the second tier. Regarding the cost allocation methods, our experiments showed diverging interests of the LSPs. While for larger LSPs, a cost allocation based on methods that lead to similar relative cost savings is advantageous, for smaller LSPs, the Shapley Value is advantageous because it pays particular attention to the very low marginal costs. Simple proportional methods based on demand characteristics such as the number or the volume of demand also lead to very different cost savings for LSPs depending on their demand characteristic. Within a coalition, it may be difficult to agree on a simple proportional method, as depending on the choice of the reference value, some LSPs will be advantaged and others disadvantaged.

For future research, it could be exciting to investigate other aspects of cooperation, such as risk sharing, which could lead to the development of a more resilient cooperative city logistics system. With regard to the strategic level, the cooperative selection and location of sites for satellites and CDCs could offer scope for further research.

Acknowledgments

The paper was written as part of a project funded by the German Research Foundation. We would like to express our gratitude to the German Research Foundation (project number: ROCOCO FO1468/1-1) for the financial support of our research project. Additionally, our gratitude goes to the Bavarian Research Alliance for their support of the research stay in Montreal, which contributed significantly to the development of this paper. While working on this paper, the second author was Adjunct Professor, Department of Computer Science and Operations Research, Université de Montréal, Canada. We gratefully acknowledge the financial support provided by the Canadian Natural Sciences and Engineering Research Council (NSERC) through its Discovery grant program, as well as by the Fonds de recherche du Québec through their Teams and CIRRELT infrastructure grants.

References

R. Agarwal and Ö. Ergun. Network design and allocation mechanisms for carrier alliances in liner shipping. *Operations Research*, 58(6):1726–1742, 2010.

M. E. Bruni, S. Khodaparasti, and G. Perboli. A bi-level approach for last-mile delivery with multiple satellites. *Transportation Research Part C: Emerging Technologies*, 160: 104495, 2024.

E. M. Cepolina and A. Farina. A new urban freight distribution scheme and an optimization methodology for reducing its overall cost. *European Transport Research Review*, 7(1):1–14, 2015.

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

T. G. Crainic and K. H. Kim. Intermodal transportation. *Handbooks in Operations Research and Management Science*, 14:467–537, 2007.

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

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

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

T. G. Crainic, F. Errico, W. Rei, and N. Ricciardi. Modeling demand uncertainty in two-tier city logistics tactical planning. *Transportation Science*, 50(2):559–578, 2016.

T. G. Crainic, M. Gendreau, and L. Jemai. Planning hyperconnected, urban logistics systems. *Transportation Research Procedia*, 47:35–42, 2020.

T. G. Crainic, G. Perboli, and N. Ricciardi. City logistics. In *Network Design with Applications to Transportation and Logistics*, pages 505–534. Springer, 2021.

F. Cruijssen, M. Cools, and W. Dullaert. Horizontal cooperation in logistics: Opportunities and impediments. *Transportation Research Part E: Logistics and Transportation Review*, 43(2):129–142, 2007a.

F. Cruijssen, W. Dullaert, and H. Fleuren. Horizontal cooperation in transport and logistics: A literature review. *Transportation Journal*, 46(3):22–39, 2007b.

EU Commission. Guidelines on the applicability of Article 81 of the EC treaty to horizontal cooperation agreements. *Official Journal of the European Communities*, 100: 2–30, 2001.

P. Fontaine, T. G. Crainic, O. Jabali, and W. Rei. Scheduled service network design with resource management for two-tier multimodal city logistics. *European Journal of Operational Research*, 294(2):558–570, 2021.

P. Fontaine, S. Minner, and M. Schiffer. Smart and sustainable city logistics: Design, consolidation, and regulation. *European Journal of Operational Research*, 307(3):1071–1084, 2023.

M. Frisk, M. Göthe-Lundgren, K. Jörnsten, and M. Rönnqvist. Cost allocation in collaborative forest transportation. *European Journal of Operational Research*, 205(2): 448–458, 2010.

M. Gansterer and R. F. Hartl. Collaborative vehicle routing: A survey. *European Journal of Operational Research*, 268(1):1–12, 2018.

M. Gansterer, R. F. Hartl, and P. E. Salzmann. Exact solutions for the collaborative pickup and delivery problem. *Central European Journal of Operations Research*, 26: 357–371, 2018.

M. Gansterer, R. F. Hartl, and M. Savelsbergh. The value of information in auction-based carrier collaborations. *International Journal of Production Economics*, 221: 107485, 2020.

D. Gillies. Solutions to general non-zero-sum games. In *Contributions to the Theory of Games*, volume 4, pages 47–85. Princeton University Press, 1959.

M. Guajardo and M. Rönnqvist. A review on cost allocation methods in collaborative transportation. *International Transactions in Operational Research*, 23(3):371–392, 2016.

A. Kimms and I. Kozeletskyi. Core-based cost allocation in the cooperative traveling salesman problem. *European Journal of Operational Research*, 248(3):910–916, 2016a.

A. Kimms and I. Kozeletskyi. Shapley value-based cost allocation in the cooperative traveling salesman problem under rolling horizon planning. *EURO Journal on Transportation and Logistics*, 5(4):371–392, 2016b.

S. Lone, N. Harboul, and J. Weltevreden. *European E-commerce Report*, 2021.

S. Mancini, M. Gansterer, and R. F. Hartl. The collaborative consistent vehicle routing problem with workload balance. *European Journal of Operational Research*, 293(3): 955–965, 2021.

J. R. Montoya-Torres, A. Muñoz-Villamizar, and C. A. Vega-Mejía. On the impact of collaborative strategies for goods delivery in city logistics. *Production Planning & Control*, 27(6):443–455, 2016.

S. Nataraj, D. Ferone, C. Quintero-Araujo, A. Juan, and P. Festa. Consolidation centers in city logistics: A cooperative approach based on the location routing problem. *International Journal of Industrial Engineering Computations*, 10(3):393–404, 2019.

S. Pan, D. Trentesaux, E. Ballot, and G. Q. Huang. Horizontal collaborative transport: Survey of solutions and practical implementation issues. *International Journal of Production Research*, 57(15-16):5340–5361, 2019.

D. N. Parisa, K. E. Samuel, and M.-L. Espinouse. Systematic literature review on city logistics: Overview, classification and analysis. *International Journal of Production Research*, 57(3):865–887, 2019.

G. Perboli, L. Brotcorne, M. E. Bruni, and M. Rosano. A new model for last-mile delivery and satellite depots management: The impact of the on-demand economy. *Transportation Research Part E: Logistics and Transportation Review*, 145:102184, 2021.

E. Pérez-Bernabeu, A. A. Juan, J. Faulin, and B. B. Barrios. Horizontal cooperation in road transportation: A case illustrating savings in distances and greenhouse gas emissions. *International Transactions in Operational Research*, 22(3):585–606, 2015.

D. Pisinger and S. Ropke. A general heuristic for vehicle routing problems. *Computers & Operations Research*, 34(8):2403–2435, 2007.

A. Rusich et al. *Collaborative Logistics Networks*. PhD thesis, Università degli Studi di Trieste, 2017.

M. Savelsbergh and T. van Woensel. City logistics: Challenges and opportunities: 50th anniversary invited article. *Transportation Science*, 50:579–590, 2016.

D. Schmeidler. The nucleolus of a characteristic function game. *SIAM Journal on Applied Mathematics*, 17(6):1163–1170, 1969.

J. Schmidt, C. Tilk, and S. Irnich. Using public transport in a 2-echelon last-mile delivery network. *European Journal of Operational Research*, 317(3):827–840, 2022.

L. S. Shapley. A value for n-person games. In H. W. Kuhn and A. W. Tucker, editors, *Contributions to the Theory of Games II*, pages 307–317. Princeton University Press, Princeton, 1953.

B. C. Shelbourne, M. Battarra, and C. N. Potts. The vehicle routing problem with release and due dates. *INFORMS Journal on Computing*, 29(4):705–723, 2017.

N. Sluijk, A. M. Florio, J. Kinable, N. Dellaert, and T. Van Woensel. Two-echelon vehicle routing problems: A literature review. *European Journal of Operational Research*, 304 (3):865–886, 2023.

S. H. Tijs and T. S. Driessen. Game theory and cost allocation problems. *Management Science*, 32(8):1015–1028, 1986.

United Nations. *World Urbanization Prospects*. United Nations, Dept. of Economic and Social Affairs, Population Division, 2018. URL `https://population.un.org/wup/publications/`.

C. Vanovermeire, D. Vercruysse, and K. Sörensen. Analysis of different cost allocation methods in a collaborative transport setting. *International Journal of Engineering Management and Economics*, 4(2):132–150, 2014.

L. Verdonck, P. Beullens, A. Caris, K. Ramaekers, and G. K. Janssens. Analysis of collaborative savings and cost allocation techniques for the cooperative carrier facility location problem. *Journal of the Operational Research Society*, 67(6):853–871, 2016.

M. Winkenbach, P. R. Kleindorfer, and S. Spinler. Enabling urban logistics services at la poste through multi-echelon location-routing. *Transportation Science*, 50(2):520–540, 2016.

A MILP notation

Table 8: Sets, parameters and decision variables

Sets and subsets	
$\mathcal{A}$	set of arcs
$\mathcal{A}(k)$	set of arcs city freighter k can use
$\mathcal{D}$	set of demands
$\mathcal{D}(n)$	set of demands assigned to LSP n
$\delta_k^-(i)$	set of arcs for city freighter k that start in node i
$\delta_k^-(i)$	set of arcs for city freighter k that end in node i
$\mathcal{E}$	set of CDCs
$\mathcal{K}$	set of city freighters
$\mathcal{K}(n)$	set of city freighters of LSP n
$\mathcal{K}(s)$	set of city freighters available at satellite s
$\mathcal{M}$	set of modes of transportation
$\mathcal{N}$	set of all LSPs
P	set of periods
$\mathcal{R}$	set of services
$\mathcal{R}(d,s)$	set of services fulfilling time windows for demand d and satellite s
$\mathcal{R}(n)$	set of services that LSP n operates
$\mathcal{R}(p,e,t)$	set of services of type t, starting from CDC e and operating during period p
$\mathcal{R}(p,s)$	set of services operating at satellite s during period p
$\mathcal{R}(p,s,m)$	set of services of mode m operating at satellite s during period p
$\mathcal{S}$	set of satellite platforms
$\mathcal{T}$	set of urban vehicle types
$\mathcal{T}(m)$	set of urban vehicle types for mode m
Parameters	
a_{spn}	maximum number of urban vehicles that LSP n is allowed to accommodate at satellite s in period p
$\bar{a}_{spmn}$	maximum number of urban vehicles of mode m that LSP n is allowed to accommodate at satellite s in period p
b_d	latest possible delivery time period for demand d
c_r	operating cost of service r
$\hat{c}_{ij}$	cost of city freighter using arc (i,j)
e_r	CDC of urban vehicle service r
f_{de}	assignment cost of demand d to CDC e
g_{spn}	total volume of freight that LSP n is allowed to accommodate at satellite s in period p
h_{etn}	fleet size of urban vehicles at CDC e of type t owned by LSP n
l	service time of city freighters at each demand location
M_1	Big M 1 for Constraints (17)
M_2	Big M 2 for Constraints (22)
m_r	transportation mode of service r
q	capacity of city freighters
ρ	handling time for each demand at each satellite
s_d	service time at demand location d.
s_k^+	satellite at which city freighter k starts
s_k^-	satellite at which city freighter k ends
$\bar{t}_{ij}$	second-tier travel time between node i and node j
t_r	urban vehicle type of service r
τ_{rs}	arrival time of service r at satellite s
u_t	urban vehicle capacity of type t
v_d	volume of demand d
w_r	service time of service r at each satellite
Decision variables	
p_{ik}	variable specifying the start of service time at vertex i serviced by city freighter k
x_{dsr}	taking the value one if demand d is assigned to satellite s and service r, zero otherwise
y_r	taking the value one if service r is selected, zero otherwise
z_{ijk}	taking the value one if city freighter k uses arc $(i,j) \in \mathcal{A}$, zero otherwise

B Performance benchmark

Table 9: Comparison of Gurobi and the I2S-LNS with varying $|D|$ and constant $|R| = 24$

Instance*	Costs	GAP [%]	Time [s]	Best costs	Avg. costs	σ [%]	Avg. time [s]	Δ[%]
N1-D5-1	94.55	0.00	1.25	94.55	94.55	0.00	10.24	0.00
N1-D5-2	77.15	0.00	0.50	77.15	77.15	0.00	17.19	0.00
N1-D5-3	121.71	0.00	4.97	121.71	121.71	0.00	19.56	0.00
N1-D5-4	123.86	0.00	2.09	123.86	123.86	0.00	17.95	0.00
N1-D5-5	126.97	0.00	1.41	126.97	126.97	0.00	18.95	0.00
N1-D10-1	173.41	0.00	209.70	173.41	173.41	0.00	48.73	0.00
N1-D10-2	177.29	4.92	3600	177.29	177.29	0.00	40.54	0.00
N1-D10-3	164.92	0.00	1045.30	164.92	164.92	0.00	31.87	0.00
N1-D10-4	166.96	0.93	3600	166.96	166.96	0.00	29.93	0.00
N1-D10-5	165.58	0.00	197.92	165.58	165.58	0.00	35.57	0.00
N1-D15-1	233.22	19.39	3600	228.43	228.96	0.23	102.20	-1.83
N1-D15-2	219.92	16.49	3600	211.45	211.67	0.10	111.50	-3.30
N1-D15-3	241.12	23.00	3600	228.08	228.08	0.00	144.20	-5.41
N1-D15-4	203.75	24.95	3600	203.75	203.75	0.00	94.96	0.00
N1-D15-5	273.66	40.92	3600	259.32	260.30	0.38	82.57	-4.52
N1-D20-1	307.00	32.73	3600	299.35	300.43	0.36	187.74	-2.15
N1-D20-2	268.18	16.76	3600	264.77	264.84	0.03	210.58	-1.25
N1-D20-3	306.51	24.92	3600	276.44	276.44	0.00	218.21	-9.80
N1-D20-4	258.46	28.23	3600	248.62	250.02	0.56	158.62	-3.26
N1-D20-5	287.34	18.48	3600	287.34	287.34	0.00	125.33	0.00
N2-D30-1	396.12	41.11	3600	342.06	343.05	0.29	273.63	-13.41
N2-D30-2	388.48	30.47	3600	382.67	386.09	0.89	486.02	-0.61
N2-D30-3	429.20	31.15	3600	386.41	390.57	1.08	297.51	-8.98
N2-D30-4	406.94	28.96	3600	370.28	371.59	0.35	500.25	-8.69
N2-D30-5	443.28	37.86	3600	376.48	377.10	0.16	467.68	-14.93
N2-D40-1	-	-	3600	487.97	491.03	0.75	600.00	-
N2-D40-2	-	-	3600	484.37	485.43	0.22	592.30	-
N2-D40-3	-	-	3600	492.01	499.20	1.46	573.82	-
N2-D40-4	-	-	3600	488.69	489.23	0.11	513.37	-
N2-D40-5	-	-	3600	480.22	485.74	0.73	442.63	-

*N1-D5-1 refers to the first out of five instances with network N1 and $|\mathcal{D}| = 5$.

Table 10: Comparison of Gurobi and the I2S-LNS with varying $|R|$ and constant $|D| = 15$

	Gurobi			I2S-LNS				
Instance*	Costs	GAP [%]	Time [s]	Best costs	Avg. costs	σ [%]	Avg. time [s]	Δ[%]
R12-1	268.74	23.39	3600	268.74	268.74	0.00	107.00	0.00
R12-2	213.72	10.20	3600	209.17	209.17	0.00	63.50	-2.13
R12-3	234.14	12.53	3600	224.77	224.77	0.00	85.12	-4.00
R12-4	227.22	16.53	3600	224.20	224.20	0.00	106.96	-1.33
R12-5	206.22	13.40	3600	206.22	206.22	0.00	101.56	0.00
R36-1	203.11	14.76	3600	202.34	202.35	0.00	80.92	-0.37
R36-2	205.21	13.57	3600	205.21	205.21	0.00	95.83	0.00
R36-3	258.75	26.47	3600	236.47	236.48	0.00	165.57	-8.61
R36-4	201.09	13.88	3600	201.09	201.09	0.00	77.94	0.00
R36-5	245.40	34.10	3600	232.16	232.16	0.00	99.09	-5.40
R60-1	206.49	22.48	3600	197.38	197.50	0.06	96.84	-4.35
R60-2	198.17	21.86	3600	191.14	191.14	0.00	110.29	-3.55
R60-3	200.71	17.36	3600	193.26	193.36	0.05	89.83	-3.66
R60-4	179.82	31.15	3600	179.82	179.82	0.00	103.94	-0.00
R60-5	240.74	30.07	3600	232.75	232.75	0.00	145.33	-3.32

*R12-1 refers to the first out of five instances with $|\mathcal{R}| = 12$.

Table 11: Performance of the I2S-LNS on larger instances

$\lvert D\rvert$	$\lvert R\rvert$	Best costs	Avg. costs	σ [%]	Avg. time [s]
50	36	552.32	555.85	0.64	1283.70
50	36	559.20	563.13	0.70	1258.43
50	36	613.62	615.01	0.23	1434.71
50	36	626.99	629.19	0.35	1230.52
50	36	583.67	589.24	0.95	1134.03
75	48	895.75	904.07	0.93	1800
75	48	844.22	848.36	0.49	1800
75	48	873.83	879.79	0.68	1800
75	48	894.47	898.29	0.43	1740.36
75	48	880.42	889.61	1.04	1751.60
100	60	1105.81	1111.72	0.53	1800
100	60	1110.46	1120.61	0.91	1800
100	60	1140.63	1147.24	0.58	1800
100	60	1092.63	1099.54	0.63	1800
100	60	1077.81	1089.62	1.10	1800