Service Network Design Models for Two-tier City Logistics

Teodor Gabriel Crainic
Dept. management et technologie
École des sciences de la gestion
and
Centre Interuniversitaire de Recherche sur les
Réseaux d'Entreprise, la Logistique et le Transport (CIRRELT)
Université du Québec à Montréal

Antonino Sgalambro
Dipartimento di Statistica, Probabilità e Statistiche Applicate
Sapienza Università di Roma

Abstract

This paper focuses on two-tier City Logistics systems for advanced management of urban freight activities and, in particular, on the first layer of such systems where freight is moved from distributions centers on the outskirts of the city to satellite platforms‘by urban vehicles, from where it will be distributed to customers by other, dedicated vehicles. We address the issue of planning the services of this first tier system, that is, select services, their routes and schedules, and determine the itineraries of the customer-demand flows through these facilities and services. We propose a general scheduled service network design modeling framework that captures the fundamental concepts related to the definition of urban-vehicle tactical plans within a two-tier distribution network. We examine several operational assumptions regarding the management of the urban-vehicle fleet, the flexibility associated with the delivery of goods, and the impact of the freight transfer operations at satellites, and show how the proposed modeling framework can evolve to represent an increasing level of detail. A discussion of algorithmic perspectives completes the paper.

Key words : City Logistics, scheduled service network design, urban freight transportation, fixed charge multicommodity network design, asset management

1 Introduction

Operational excellence and efficiency in the management of freight transportation represent crucial enabling factors in the current globalized business models, aiming at the reduction of costs in shipping and handling goods, raw materials, and components. Elements such as an increasing amount of e-commerce transactions and the presence of firms within the city boundaries require the definition of improved practices and management models, designed specifically for the shape and needs of urban logistics networks. Indeed, several additional challenges rise up in the evaluation and planning of city logistics transportation systems, due to the high variability in the size and shape of roads, the presence of logistics infrastructures, and the population density on different zones of the considered area. Within the urban context, particular attention is dedicated to the issue of sustainability, taking into account different perspectives in the optimization of freight logistics operations, such as social and economic efficiency, reduction of polluting emissions and road capacity occupation, equity in the exposure of the population to the risk related to freight logistics (e.g., hazardous material transportation), and so on.

The need for such an approach yielded the idea of considering all the stakeholders and movements that take place into the urban area as parts of a single *integrated logistics system*. In this way, the consolidation of goods related to different customer-demands into the same shipment, together with the coordination of carriers and shippers, become basic requirements for a *city logistics* approach, aiming at the optimization of the whole integrated system.

Decision and management policies are commonly classified according to different planning levels: long term or strategic decisions, mid term or tactical planning, and short-term or operational level. City logistics optimization challenges involve all these planning levels. Some examples of strategic planning issues for the design of a city logistics system fall in the class of location and layout of logistics platforms, together with the selection of corridors and streets open to the different class of vehicles, and the dimensioning of the fleet. Tactical planning challenges are commonly focused on the need for consolidation processes, aiming to build efficient transportation plans taking concurrently into account the quality of the delivery service and the variability of the demand. Classical operational planning issues in city logistics concern detailed resource and vehicle scheduling, and operations management at the logistic platforms.

Most current city logistics projects involve single tier schemes, where delivery routes start directly from a single logistics platform (also called City Distribution Center, or CDC) devoted to the consolidation of freight from long-haul transportation trucks into smaller vehicles. Single tier models are sometimes extended to the case of multiple CDCs, the city area being usually partitioned in zones, each zone being assigned to a specific CDC. Single tier approaches do not appear very successful for large cities characterized by important center areas with high density of population and intense

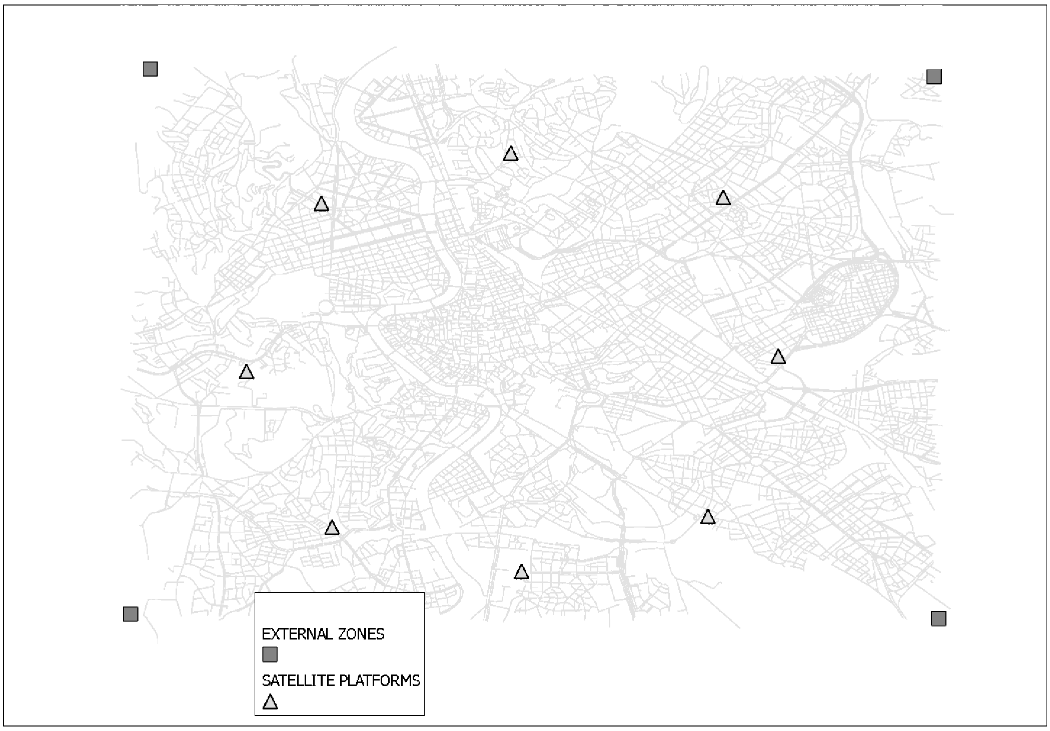

Figure 1: External zones and satellite platforms on an urban network

commercial, administrative, and cultural activities. More articulated two-tier systems were therefore proposed, such as those described in [6], which are the basis for the work reported in this paper.

A two-tier city logistics system is based on a *double layer* generalization of the City Distribution Center scheme. The presence of CDCs located at the outskirts of the urban zone permits a first level of consolidation and coordination activities for long-haul transportation vehicles. External CDCs are also referred to as *external zones*. Freight at the external zones is consolidated into urban vehicles and shipped toward a second set of infrastructures named *satellite platforms*, located inside the city (see Figure 1 for a sketch of such a two-tier structure). Urban vehicles are supposed to travel from external zones to satellites along properly selected corridors, in order to minimize the nuisance effects on urban traffic. satellites are dedicated to the transfer of freight from urban vehicles to smaller ones, also named *city-freighters*, that can travel along any street in the city center area and deliver loads to their designated consignees (final customers).

This two-tier scheme requires more complex evaluation and planning models and procedures than those developed for single tier schemes. Indeed, some of the main optimization and modeling challenges are related in this case not only to the operations of vehicles in each layer, but also to the interactions arising between the two layers of the logistics network, such as the synchronization and coordination of the fleet and terminal operations. In this context, tactical planning processes are to be studied for both evaluation and planning purposes.

A comprehensive mathematical model is provided in [6] for the problem of defining short-term tactical and operational plans for both urban-vehicle and city-freighter fleets. The resulting optimization problem, referred to as the *Day-Before Planning problem*, turns out to be the $\mathcal{NP}$-hard *two-echelon, synchronized, scheduled, multi-depot, multiple-tour, heterogeneous vehicle routing problem with time windows*. A possible way to cope with such a complex problem consists in decomposing it hierarchically into two easier subproblems, namely, the Urban-Vehicle Service Network Design problem, and the City-Freighter Circulation problem. The scope of this paper is to provide a deeper modeling and algorithmic insight for the former problem, namely, to propose a general modeling framework expressing the basic concepts related to the definition of *urban-vehicle* tactical plans within a two-tier distribution urban network. We examine several operational assumptions regarding the management of the urban-vehicle fleet, namely, the flexibility associated with the delivery of goods, and the impact of the freight transfer operations at satellites, and show how the proposed modeling framework can evolve to represent such an increasing level of detail. We also identify promising algorithmic perspectives and strategies.

The paper is organized as follows. The Urban-Vehicle Service Network Design problem is described in Section 2, and several possible variants of the basic problem, according to different levels of control and detail in urban-vehicle fleet management, are discussed. The basic version of the problem is formally described and modeled in Section 3, by introducing a time-space network representation of the problem and a fixed charge network design formulation. Sections 4, 5, and 6 present extensions of the basic model according to different management policies. Finally, a nuymber of algorithmic perspectives are discussed in Section 7.

2 Problem Description and Basic Variants

A detailed description of the considered problem requires some basic definitions, that will be used in the remainder of the paper:

Service leg An urban-vehicle *service leg* is defined as the movement performed by a capacitated urban vehicle from one facility (an external zone or a satellite) to a different one in the first tier of a City Logistic network.

Service An urban-vehicle *service* is identified by a non-empty set of consecutive urban-vehicle service legs (a route) starting from an external zone and touching a strictly positive number of satellites.

Service+ An urban-vehicle *service+* is defined by an urban-vehicle service plus the associated repositioning path toward an external zone or the vehicle depot.

Work assignment An urban-vehicle *work assignment* is defined as a sequence of services+ associated to an operating urban vehicle, according to a certain plan. Depending on the operational setup, the work assignment of an operating urban vehicle can be composed by one or more services+.

Schedule The schedule of an urban-vehicle service is defined by a *departure time* from an external zone, a route consisting in an ordered sequence of satellites to be visited, plus a departure time from each of the visited satellites that takes into account the time needed for the freight transfer operations.

The *Urban-Vehicle Service Network Design* problem (UVSND) consists in the definition of a complete schedule for the work assignment of the urban vehicles involved in the transportation from external zones to satellites of the freight associated with a given set of customer demands in a two-tier city logistics system. Urban vehicles are allowed to hold at satellites only for the time strictly needed for unloading and transfer operations. The total quantity of goods related to each customer-demand is assumed to be shipped on a single urban vehicle. Additional issues related to the design and operation of this system are discussed in the following three sub-sections.

2.1 Control, management and route restrictions for urban-vehicle services

Due to the high variability in ownership and operations of urban-vehicles, several policies can be chosen, involving different levels of control and detail in urban-vehicle fleet management. Two major variants are introduced to properly represent the possible system organization and urban-vehicle planning process.

The *uncontrolled urban-vehicle fleet* variant of the problem considers the situation in which the number of available urban vehicles is not considered explicitly, and the routes of the vehicles providing the services are not controlled once goods are delivered at satellites. The type and the load capacity of urban vehicles impacts the development of urban-vehicle operational plans, but no additional hypotheses are made concerning any central control for the fleet of vehicles,or restrictions in the number of available urban vehicles and their movements outside the controlled city logistics area. In this variant of the problem, the repositioning of urban vehicles from satellite platforms to external zones can be disregarded.

A different version of the problem, the *controlled and limited* variant, is dedicated to situations in which a higher level of control is imposed on the vehicle movements. One assumes that each vehicle starts its trip from an external zone and finishes its service at a, possibly different, external zone.

The controlled-fleet assumption thus requires the inclusion into the service network model of repositioning movements from satellites to external zones for termination of service or reloading, as well as the chance for the vehicles to wait at the external zones. The selection of services is thus accompanied by the construction of routes for the corresponding vehicles. Moreover, the presence of a fleet capacity is explicitly considered in the model, together with a limit on the maximum length for a vehicle route.

2.2 Space and time characterization of the demand at the satellites

A fundamental assumption for the problem of planning the distribution service in a two-tier logistics system consists in the presence of a given deterministic set of customer-demands. Shipping requests are described through the associated external zone and time instant the goods become available and ready for loading and transport, the final destination, as well as the type and the quantity of the products to be shipped. The set of shipping requests can be further described with an increasing number of details. We consider two different variants for the service network model, referring to the characterization of customer-demands.

Fixed delivery satellites and delivery time variant. In this variant of the problem, the set of customer locations in the network is partitioned in such a way to allocate each final customer node to a single satellite platform. Customer-demand deliveries may therefore be directly associated with a satellite, and the time interval of each delivery is strictly imposed as well. These assumptions induce a strong simplification in the optimization process, and are actually valid only in those contexts in which

- the assignment of customers to satellite platforms can be really assumed to be straightforward due to specific network related characteristics;
- the dimension of the discrete time intervals in the time horizon is large enough to represent a reasonable time window for the deliveries of customer-demands.

In more general cases, the optimization variants arising from these restrictive assumptions can still be useful to simplify real cases and derive easiest problems that can be solved in reduced computational times and, thus, provide useful bounds for the more realistic variants.

Floating delivery satellites and flexible time windows variant. There is no fixed association of final customers to satellites in this case. The problem consists in deciding the sequence of satellites for urban-vehicle services in order to serve all the customer demands, assuming that the overall distribution cost is to be minimized. Approximations of the shipping costs from satellites to the final customers are required, delivery time windows being specifically considered for each customer demand (these could cover several periods of the discrete time horizon).

2.3 Unloading operations at satellites

An important part of the overall time required by freight shipping services is devoted to unloading and transferring goods from urban vehicles to city freighters at satellites. Unloading times are assumed here to be integer multiples of the time periods, and the presence of unloading times is considered in the model with different levels of detail, as described in the following.

Fixed unloading time variant. In a first variant of the problem, we consider fixed unloading times for each urban-vehicle service. Unloading times are assumed in this case to be related to the maximum load capacity of each vehicle and to the class of product. Fixed unloading times are thus supposed to represent an upper bound on the actual amount of time urban vehicles are expected to wait at satellites in order to perform unloading operations.

Load-dependent unloading time variant. The level of detail related to the representation of urban-vehicle unloading times at satellites can be increased, with respect to the previous case, by considering load-dependent unloading times for each vehicle. According to this more realistic assumption, each vehicle waits at the satellites for an integer number of time intervals that depends on the size of the load that has to be actually discharged . This additional feature leads to a more complex representation of the planning problem, discussed in Section 6.

3 The uncontrolled urban-vehicle service network design

In this section, we consider the optimization problem arising from the combination of the basic variants of the Urban-Vehicle Service Network Design above described, namely the

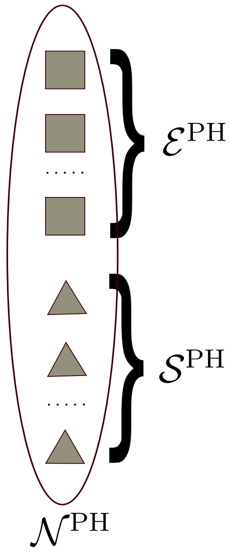

Figure 2: A scheme of the physical node set $\mathcal{N}^{\mathrm{PH}} = \mathcal{E}^{\mathrm{PH}} \cup \mathcal{S}^{\mathrm{PH}}$ as the union of external zones and satellite nodes

case with uncontrolled urban-vehicle fleet and fixed satellite allocations and delivery times for customers, and fixed unloading times at satellites. We provide a formal definition, the notation, and a mixed integer programming formulation.

3.1 Problem definition and notation

The underlying static scheme of this two-tier problem is composed of two sets of facilities: a set $\mathcal{E}^{\mathrm{PH}} = \{e\}$ of external zones and a set $\mathcal{S}^{\mathrm{PH}} = \{s\}$ of satellites, as illustrated in Figure 2. The union of these two sets gives rise to the set $\mathcal{N}^{\mathrm{PH}} = \mathcal{E}^{\mathrm{PH}} \cup \mathcal{S}^{\mathrm{PH}}$ of physical nodes of the network. Assuming the presence of a set $\mathcal{A}^{\mathrm{PH}}$ of arcs representing the roads linking these facilities, we can define the physical network $\mathcal{G}^{\mathrm{PH}} = (\mathcal{N}^{\mathrm{PH}}, \mathcal{A}^{\mathrm{PH}})$.

A time-expanded network is defined to represent the time-dependent characteristics of the problem, by replicating the sets of physical nodes $\mathcal{N}^{\mathrm{PH}}$ for each period of a finite and discrete time-horizon $\mathcal{T} = \{1, ..., T_{max}\}$. The notation used to represent the elements in the time-space network is then the following. We refer to $\mathcal{N}$ as the set of all the nodes in the time-expanded graph, $\mathcal{N} = \bigcup_{t\in\mathcal{T}} \mathcal{N}_t$, where $\mathcal{N}_t$ is the realization of the physical nodes at a certain time $t \in \mathcal{T}$. We denote by $\mathcal{E}_t$ the realization of the set of external zones $\mathcal{E}^{\mathrm{PH}}$ at time $t \in \mathcal{T}$ and by $\mathcal{S}_t$ that of the satellites $\mathcal{S}^{\mathrm{PH}}$ at the same time instant, in such a way that $\mathcal{N}_t = \mathcal{E}_t \bigcup \mathcal{S}_t$, $\mathcal{E} = \bigcup_{t\in\mathcal{T}} \mathcal{E}_t$ and $\mathcal{S} = \bigcup_{t\in\mathcal{T}} \mathcal{S}_t$. We mark $t(i)$ the instant in the time horizon associated to node i in the time-space representation. A sketch of the time-expanded set of nodes is depicted in Figure 3.

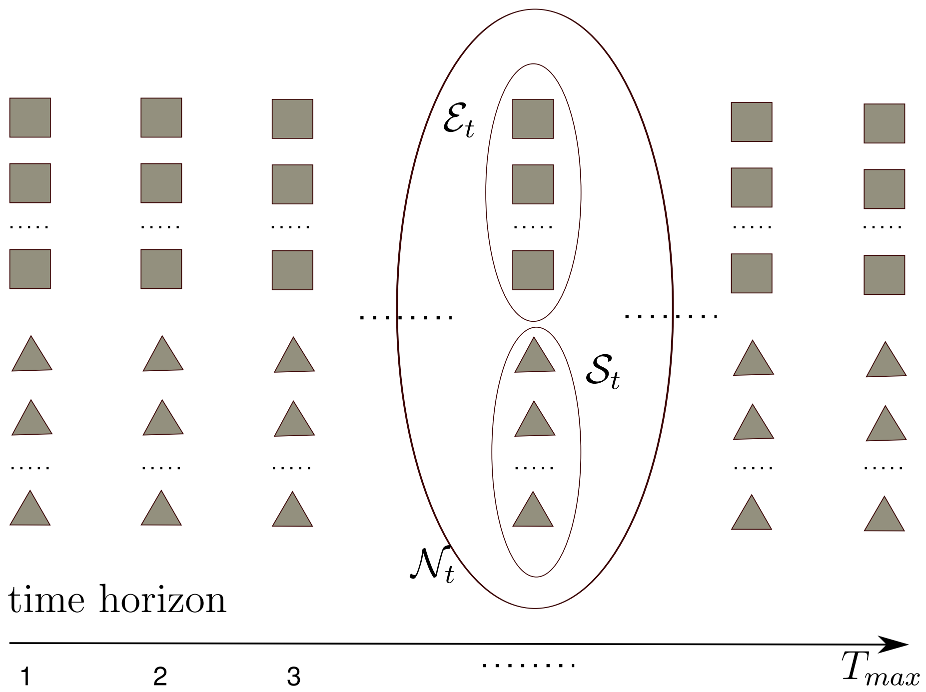

Figure 3: Representation of the complete set of nodes $\mathcal{N}$ of the time-space network over the time horizon $\mathcal{T} = \{1, ..., T_{max}\}$.

The set of arcs $\mathcal{A}$ of the time-space network $\mathcal{G} = (\mathcal{N}, \mathcal{A})$ is composed by three subsets of arcs, namely the *movement arcs* $\mathcal{A}^{\mathrm{M}}$ that connect nodes representing different physical nodes in the time-space network, the *holding arcs* $\mathcal{A}^{\mathrm{H}}$ that link couples of nodes representing the same external zone at different time periods, and the *unloading arcs* $\mathcal{A}^{\mathrm{U}}$ connecting the nodes representing the same satellites at different time periods. The holding arcs express the waiting times spent by the urban vehicles at the external zones. We assume it is not possible to hold vehicles at the satellite nodes, except for the time strictly needed to perform the unloading operations, represented by the unloading arcs. For each node i, we define the set $\mathcal{N}^+(i) = \{j \in \mathcal{N} : (i, j) \in \mathcal{A}\}$ of successor nodes and the set $\mathcal{N}^-(i) = \{j \in \mathcal{N} : (j, i) \in \mathcal{A}\}$ of predecessor nodes.

Define the customer demands $d \in \mathcal{D}$ characterized by a quantity $w(d)$, an origin external zone node $e(d) \in \mathcal{E}$, and a destination node $s(d) \in \mathcal{S}$. Given the hypotheses assuming fixed satellite allocations and delivery times for customer demands, each shipping request is perfectly identified in space and time by the couple of nodes $(e(d), s(d))$ in the time-space network. The presence of different classes of products $p \in \mathcal{P}$ is assumed, the class of product of customer-demand d being denoted $p(d)$.

Our goal is to propose a model that will yield an "optimal" service plan and schedule for urban vehicles. The scheduling requirement has led to the construction of the time-space network defined above. According to the definitions of Section 2, the arcs of set $\mathcal{A}^{\mathrm{M}}$

are service legs and one must determine the urban-vehicle services and schedules, that is, the sequence of satellites visited and the time departure from the external zone for each urban vehicle. In an arc-based formulation, however, it is not necessary to define *a priori* all the potential service routes. It is sufficient to identify the time departure from the external zone and let the model select the services to be provided and the route of each vehicle. In the rest of this section, we will therefore refer interchangeably to service departure and service.

Define a set $\mathcal{R} = \{r\}$ of possible urban-vehicle service departures, with cardinality $|\mathcal{R}|$. Different classes of products $p \in \mathcal{P}$ could require particular types of urban vehicles (for instance perishable products, hazardous materials, etc.). Therefore, we introduce several subsets of urban-vehicle service departures $\mathcal{R}^p, p \in \mathcal{P}$, each one associated with a certain product class; $\mathcal{R} = \bigcup_{p \in P} \mathcal{R}^p$.

Three types of costs are considered in this problem: a fixed cost $f_{p(r)}$ for each urban-vehicle service departure $r \in \mathcal{R}^p$, service legs costs k_{ijr} associated with the activation of a service leg related to service r between nodes i and j, and shipping costs c^p_{ijr} associated with the transfer of a unit of product p from node i to node j on service r in the time-space network. The load capacity provided by the selection of the service departure $r \in \mathcal{R}$ is indicated with u_r, while λ_i is the maximum number of urban vehicles that can concurrently dock and be unloaded at satellite $i \in \mathcal{S}$.

We denote$\mathcal{W}_r(i) = \{j \in \mathcal{N} : (i,j) \in \mathcal{A}^{\mathrm{U}}\}$ the set of the head nodes for the unloading arcs departing from node i with respect to service r in the time-space representation. In the fixed unloading time variant of the problem, only one holding arc leaves each satellite node $i \in \mathcal{S}$, and hence $|\mathcal{W}_r(i)| = 1 \forall i, r$. A more complex notation for the load-dependent case is presented in Section 4.

3.2 A Fixed Charge Network Design mathematical formulation

Service Network Design problems are generally formulated in the literature as Fixed Charge Network Design problems (see for instance [4]). In particular, for the tactical planning of consolidation-based carriers, multicommodity capacitated optimization problems were recently proposed in [1] and [10]. Following the same modeling approach, sets of decision variables must be introduced to represent the decisions represented in the formulaiton. In particular, for each possible urban-vehicle service departure $r \in \mathcal{R}$ we introduce:

- a binary variable ϕ_r assuming a value equal to 1 if the service r is actually selected, and 0 otherwise;
- a set of service design variables $y_{ijr}, (i,j) \in \mathcal{A}$, defining the service legs associated

to service r: y_{ijr} assumes a value equal to 1 if the arc (i,j) is included in service r, and 0 otherwise;

- a set of flow variables $x^d_{ijr}, (i,j) \in \mathcal{A}, d \in \mathcal{D}$, representing the amount of goods of customer-demand d carried by the service r along the arc (i,j);
- a set of binary variables $z^d_r, d \in \mathcal{D}$, assuming a value equal to 1 if the goods related to the customer-demand d are shipped through the service r, and 0 otherwise.

Following the classification proposed in [1], we refer to the proposed formulation as the *aa-uUVSND* model, i.e., the arc-flow/arc-design uncontrolled urban-vehicle Service Network Design model. The model is displayed in Figure 4. The objective function aims at the minimization of the sum of fixed costs, service costs, and shipping costs. Constraints (2) ensure the conservation of flows at nodes and the satisfaction of the customer-demands between external zones and satellites, together with constraints (3) assigning each customer-demand to exactly one urban-vehicle service departure. Constraints (4) activate urban-vehicle legs and impose limits on the amount of flow on each leg, while constraints (5) enable the urban-vehicle service activation variables ϕ_r and forbid, for each period in the time horizon, the presence of the same service on more than one arc. Relations (6) are introduced to force urban-vehicles to wait at the satellites for the time required to perform the unloading operations. Finally, constraints (7) are devoted to impose limits on the number of urban-vehicle services that can simultaneously interest a satellite $i \in \mathcal{S}$. Unloading contraints (6) can be equivalently expressed by:

$$z^d_r - \sum_{i \in \mathcal{W}_r(s(d))} y_{s(d)ir} = 0, d \in \mathcal{D}, r \in \mathcal{R} \qquad (12)$$

4 The controlled urban-vehicle service network design problem with fleet restrictions

The *aa-uUVSND* can be modified to address the controlled fleet variant presented in the previously. There are some features that must be considered when dealing with this variant of the problem:

1. the route associated with each operating urban-vehicle service must start and finish at external zones, that is, it must be a urban-vehicle service+;
2. there is a maximum number of available urban vehicles for each class of service $\mathcal{R}^p$, denoted by n_p;
3. a maximum length equal to L_r must be considered for each urban-vehicle service $r \in \mathcal{R}$.

$$\min \sum_{r \in \mathcal{R}} \left[f_{p(r)} \phi_r + \sum_{(i,j) \in \mathcal{A}} k_{ijr} y_{ijr} + \sum_{(i,j) \in \mathcal{A}} \sum_{d \in \mathcal{D}} c_{ij}^{p(d)} x_{ijr}^d \right] \tag{1}$$

$$\text{s.t.} \sum_{j \in \mathcal{N}^+(i)} x_{ijr}^d - \sum_{j \in \mathcal{N}^-(i)} x_{jir}^d = \begin{cases} w(d) z_r^d, & i = e(d) \\ -w(d) z_r^d, & i = s(d) \\ 0, & otherwise \end{cases} \qquad d \in \mathcal{D}, r \in \mathcal{R}^{p(d)}, i \in \mathcal{N} \tag{2}$$

$$\sum_{r \in \mathcal{R}} z_r^d = 1 \qquad d \in D \tag{3}$$

$$\sum_{d \in \mathcal{D}} x_{ijr}^d \leq y_{ijr} u_r \qquad (i,j) \in \mathcal{A}, r \in \mathcal{R} \tag{4}$$

$$\sum_{(i,j) \in \mathcal{A}:\ t(i) \leq t < t(j)} y_{ijr} - \phi_r \leq 0 \qquad t \in \mathcal{T}, r \in \mathcal{R} \tag{5}$$

$$\sum_{i \in \mathcal{N}^-(s(d))} x_{is(d)r}^d - \sum_{i \in \mathcal{W}_r(s(d))} w(d) y_{s(d)ir} = 0 \qquad d \in \mathcal{D}, r \in \mathcal{R} \tag{6}$$

$$\sum_{r \in \mathcal{R}} \sum_{j \in \mathcal{N}^-(i)} y_{jir} \leq \pi_i \qquad i \in \mathcal{S} \tag{7}$$

$$\phi_r \in \{0,1\} \qquad r \in \mathcal{R} \tag{8}$$

$$y_{ijr} \in \{0,1\} \qquad (i,j) \in \mathcal{A}, r \in \mathcal{R} \tag{9}$$

$$z_r^d \in \{0,1\} \qquad d \in \mathcal{D}, r \in \mathcal{R}^{p(d)} \tag{10}$$

$$x_{ijr}^d \geq 0 \qquad (i,j) \in \mathcal{A}, d \in \mathcal{D}, r \in \mathcal{R}^{p(d)} \tag{11}$$

Figure 4: The *aa-uUVSND* model

The first feature can be accomplished by introducing in the set $\mathcal{N}$ an additional dummy node γ connected to each external zone node $i \in \mathcal{E}$ in the time-space network by means of two dummy arcs: an arc $(i, \gamma) \in \mathcal{A}$ with associated service costs $k_{i\gamma r} = 0, \forall r \in \mathcal{R}$, and an arc $(\gamma, i) \in \mathcal{A}$ with associated service costs $k_{\gamma ir} = 0, \forall r \in \mathcal{R}$. In order to control the complete route of the urban vehicles, defining a single unsplit circular route for each of the operating urban vehicle, starting from node γ, passing through an external zone, a set of satellites, and then again through an external zone before getting back to γ, the following sets of constraints are added:

$$\sum_{j \in \mathcal{N}^+(i)} y_{ijr} - \sum_{j \in \mathcal{N}^-(i)} y_{jir} = 0 \qquad r \in \mathcal{R}, i \in \mathcal{N} \tag{13}$$

$$\sum_{i \in \mathcal{N}^+(\gamma)} y_{\gamma ir} - \phi_r = 0 \qquad r \in \mathcal{R} \tag{14}$$

The second feature, enforcing the maximum number of available vehicles for each class of urban-vehicle services, can be addressed by adding to the formulation the following set of constraints:

$$\sum_{r \in \mathcal{R}^p} \phi_r \leq n_p \qquad p \in \mathcal{P} \tag{15}$$

With respect to the third feature, a limit L_r on the route length for each service $r \in \mathcal{R}$ can be imposed by introducing the following set of additional constraints:

$$\sum_{(i,j) \in \mathcal{A}} l_{ij} y_{ijr} \leq L_r \qquad r \in \mathcal{R}, \tag{16}$$

where l_{ij} is the length of arc $(i, j) \in \mathcal{A}$. In the following we will refer to the complete (1)-(16) formulation as the *aa-cUVSND* model.

5 Floating delivery satellites and flexible time windows variant

It is often possible to reach a final customer by delivery routes starting from more than one satellite. Moreover, it can be worth to consider a set $[a(d), b(d)] \subset \mathcal{T}$ of feasible delivery times for the final customer, rather than the single instant as in the above presented model. Such a generalization can be enabled by selecting, for each $d \in \mathcal{D}$, a set denoted $Q(d)$ containing all the nodes in $\mathcal{S}$ that, by loading goods on a city freighter at time $t(i)$, it is possible to reach the final customer node within the given time windows $[a(d), b(d)]$. In order to complete the representation, for each $d \in \mathcal{D}$, we include in the time-expanded network:

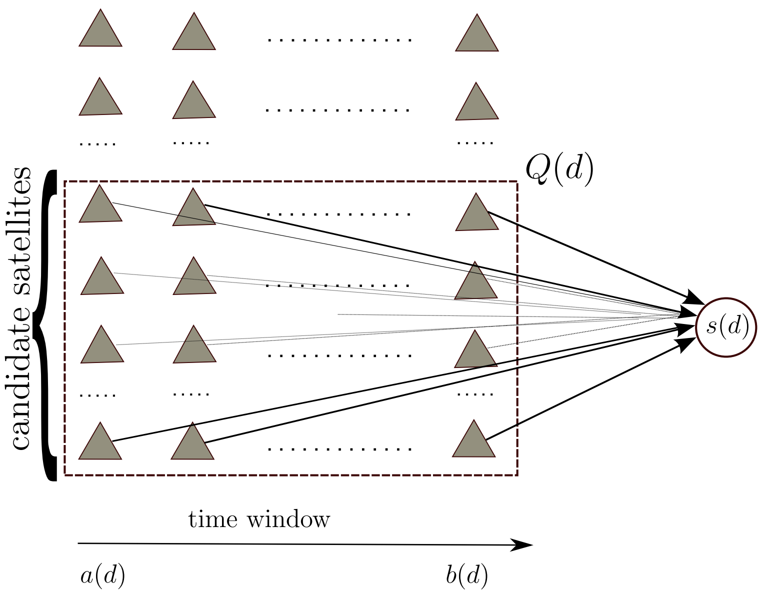

Figure 5: A scheme of the candidate satellite set and time windows subnetwork related to a customer-demand delivery node

- an additional node $s(d)$ representing the delivery of a customer-demand, $d \in \mathcal{D}$;
- a set $\mathcal{A}^d$ of additional arcs $(i, s(d))\ \forall i \in Q(d)$.

A sketch of this additional representation is displayed in Figure 5.

Only flow variables x^d_{ijr} are associated with arcs in $\mathcal{A}^d$, while the constraints referring to the service design variables do not interest these additional arcs. Exploiting this modified network representation, constraints (2) are now consistent with the presence of several candidate satellites and a time windows for the delivery of each customer-demand. Note that, by construction, the set $Q(d)$ coincides with the set $\mathcal{N}^-(s(d))$. This generalization allows the model to include the limits concerning the maximum number of city-freighter vehicles that can be loaded concurrently at a certain satellite. This can be modeled by introducing the following set of additional restrictions:

$$\sum_{d \in \mathcal{D}} \sum_{r \in \mathcal{R}^{p(d)}} \sum_{j \in \mathcal{N}^-(i)} x^d_{ijr} \leq \lambda_i u_s i \in \mathcal{S}, \tag{17}$$

where u_s represents the capacity of a city-freighter and λ_i is the maximum number of these vehicles that can be loaded concurrently at satellite i, $i \in S$.

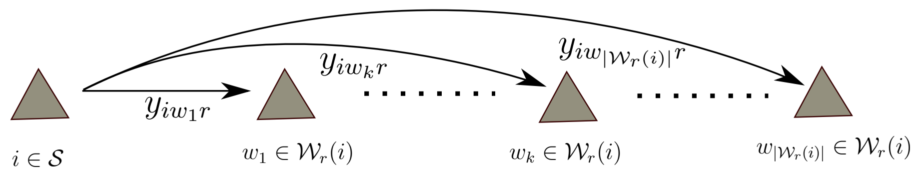

Figure 6: Sketch of the unloading arcs departing from a satellite node $i \in \mathcal{S}$ with respect to a certain urban-vehicle service $r \in \mathcal{R}$ showing the associated service design variables

6 Load-dependent unloading time model

In order to generalize the model to the representation of load-dependent unloading times, we recall the definition of $\mathcal{W}_r(i) = \{j \in \mathcal{N} : (i,j) \in \mathcal{A}^{\mathrm{U}}\}$ as the set of head nodes related to the unloading arcs departing from node i for service r. We generalize now the use of this notation to the case in which a node i and an urban-vehicle service r give rise to a certain number $|\mathcal{W}_r(i)| \geq 1$ of unloading arcs in the time-space network. We assume w.l.o.g. that the head nodes in $\mathcal{W}_r(i)$ are ranked in such a way that $t(k) > t(k-1) \forall k \in \mathcal{W}_r(i)$, and we adopt the notation w_k to refer to the holding node position k in the set $W_r(i)$. The set of unloading arcs gives rise to an equal number of *unloading classes*. Unloading classes are defined starting from the maximum load capacity u_r of each vehicle $r \in \mathcal{R}$: the k-th class will be associated with a quantity φ of flow to be discharged such that:

$$\frac{(k-1)u_r}{|\mathcal{W}_r(i)|} < \varphi \leq \frac{ku_r}{|\mathcal{W}_r(i)|} k \in \{1, ..., |\mathcal{W}_r(i)|\} \tag{18}$$

The load level of an urban vehicle that unloads at a satellite will fall therefore in exactly one out of the $|\mathcal{W}_r(i)|$ unloading classes. The load-dependent unload time model is then obtained by adding the following sets of constraints to the formulation:

$$M(1 - y_{iw_kr}) \geq \varphi_r(i) - ku_r/|\mathcal{W}_r(i)| \qquad i \in \mathcal{S}, r \in \mathcal{R}, k \in \{1, ..., |\mathcal{W}_r(i)|\} \tag{19}$$

$$M(1 - y_{iw_kr}) \geq (k-1)u_r/|\mathcal{W}_r(i)| - \varphi_r(i) + \epsilon \quad i \in \mathcal{S}, r \in \mathcal{R}, k \in \{1, ..., |\mathcal{W}_r(i)|\}, \tag{20}$$

where

$$\varphi_r(i) = \sum_{d \in \mathcal{D}: i = s(d)} \sum_{j \in \mathcal{N}^-(i)} x^d_{ijr} = \sum_{d \in \mathcal{D}: i = s(d)} z^d_r w(d)$$

and the index iw_kr stands for the k-th element in the ordered set of holding arcs going out from each node i and vehicle r, as illustrated in Figure 6. Constraints (19) and (20) are used to activate only the feasible class according to the quantity of products to be discharged, ϵ being a small positive quantity introduced to break possible ties among classes.

7 Algorithmic perspectives

The models proposed in this paper for the Urban-Vehicle Service Network Design problem are based on the widely studied class of capacitated multicommodity fixed charge network design problems (CMND). A review on the recent literature on algorithms for the CMND problem can hence reveal promising approaches for coping with all the variants of the considered problem. We start by recalling that realistically sized instances of the fixed-charge network design problem are rarely solved to the optimum by pure exact methods. A cutting-plane algorithm was recently proposed in [2] embedding efficient separation and lifting procedures for several families of valid inequalities, such as strong, cover, minimum cardinality, flow cover and flow pack inequalities. The experimental results of this approach confirmed the effectiveness of the considered sets of inequalities (in particular cover and flow cover inequalities) to improve the results in comparison with state-of-art solvers.

Benders decomposition approaches were proposed as well for this class of network design problems. A review on the application of this method can be found in [3]. Several interesting heuristics methods were also proposed in the literature for this problem. In [5], a Lagrangian relaxation approach was presented, while in [7] a class of neighborhoods for meta-heuristic algorithms were proposed. The basic idea in this paper was to explore the space of the arc design variables by re-directing flow around cycles and closing and opening design arcs accordingly. In this way, the neighborhood can take into account the impact on the total design cost of potential modifications to the flow distribution of several commodities concurrently. Such a scheme was embedded in a tabu-search algorithm and produced high quality experimental results, outperforming those of other heuristics previously presented in the literature.

More recent heuristic approaches rely predominantly on *math-heuristics*, that is, methods based on a combined use of meta-heuristic schemes and mathematical programming. A local branching meta-heuristic based on the use of a general MIP solver to explore the neighborhood was proposed in [11] with promising results. In [8], a fast local search heuristic based on IP exploration of the neighborhood was studied, exploiting both compact arc-based and extended path-based formulation for the considered problem. The method produced good quality solutions in quick computational times and can be easily applied on different variants of the problem. A capacity scaling heuristic was introduced in [9] by combining a column and row generation technique and a strong formulation including forcing constraints.

The application of math-heuristic methods for the proposed arc-based formulation of the UVSND problem seems to be a promising way to address it. Indeed, the results of such methods are enhanced by the availability of bounds provided by good quality feasible solutions. For the considered problem, several approaches can be followed to obtain such solutions on given instances of the problem. Given a physical network and a set of customer-demands, and depending of the required level of approximation, it is always possible to solve a reduced version of the problem arising from the exploitation of some of the presented basic variants. For instance, the size of the time-space network can be reduced by considering aggregated time-periods and exploiting the solution obtained on these simplified instances as an upper bound for the original problem. Similarly, floating delivery satellites can be temporarily fixed by assigning customers to the closest satellite: thus, solving the fixed satellite variant of the problem could permit to find a feasible solution on a reduced size network.

With respect to the *aa-cUVSND* formulation presented in Section 4 and related to the controlled variant of the problem, the presence of the additional set of constraints (13), known in the literature as design-balance constraints, was treated in some recent scientific contributions in the field of service network design. In particular, in [10], a tabu search procedure was proposed for the design-balanced variant of the capacitated multicommodity network design problem, based on a two local search phases exploiting an hybrid of an add/drop procedure and single-path-based neighborhood. In [1] arc-based formulations for design-balanced SND problems were compared with path-based and cycle-based formulations, showing empirically how the latter yields to better computational performances. This aspect should be hence considered in the design of algorithms for the controlled variant of the UVSND, by coupling the *aa-cUVSND* formulation with a cycle-based reformulation of the problem.

Acknowledgments

While working on this project, the first author was the NSERC Industrial Research Chair on Logistics Management, ESG UQAM, and Adjunct Professor with the Department of Computer Science and Operations Research, Université de Montréal, and the Department of Economics and Business Administration, Molde University College, Norway. Partial funding for this project has been provided by the Natural Sciences and Engineering Council of Canada (NSERC), through its Industrial Research Chair and Discovery Grants programs, and by the partners of the Chair, CN, Rona, Alimentation Couche-Tard and the Ministry of Transportation of Québec.

References

[1] J. Andersen, T. G. Crainic, and M. Christiansen. Service network design with asset management: Formulations and comparative analyses. *Transportation Research Part C*, 17:197–207, 2009.

[2] M. Chouman, T. G. Crainic, and B. Gendron. A cutting-plane algorithm for multicommodity capacitated fixed-charge network design. Technical report, CIRRELT, 2009. CIRRELT-2009-20.

[3] A. M. Costa. A survey on benders decomposition applied to fixed-charge network design problems. *Comput. Oper. Res.*, 32(6):1429–1450, 2005.

[4] T. G. Crainic. Service network design in freight tranportation. *European Journal of Operational Research*, 122:272–288, 2000.

[5] T. G. Crainic, A. Frangioni, and B. Gendron. Bundle-based relaxation methods for multicommodity capacitated fixed charge network design. *Discrete Appl. Math.*, 112(1-3):73–99, 2001.

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

[7] I. Ghamlouche, T. G. Crainic, and M. Gendreau. Cycle-Based Neighbourhoods for Fixed-Charge Capacitated Multicommodity Network Design. *OPERATIONS RESEARCH*, 51(4):655–667, 2003.

[8] M. Hewitt, G. L. Nemhauser, and M. W. P. Savelsbergh. Combining Exact and Heuristic Approaches for the Capacitated Fixed-Charge Network Flow Problem. *INFORMS JOURNAL ON COMPUTING*, page ijoc.1090.0348, 2009.

[9] N. Katayama, M. Chen, and M. Kubo. A capacity scaling heuristic for the multicommodity capacitated network design problem. *J. Comput. Appl. Math.*, 232(1):90–101, 2009.

[10] M. B. Pedersen, T. G. Crainic, and O. B. G. Madsen. Models and tabu search metaheuristics for service network design with asset-balance requirements. *Transportation Science*, 43(2):158–177, 2009.

[11] I. Rodríguez-Martín and J. José Salazar-González. A local branching heuristic for the capacitated fixed-charge network design problem. *Comput. Oper. Res.*, 37(3):575–581, 2010.