CIRRELT-2024-31

Service Network Design for Consolidation-based Transportation – Advanced Topics

Teodor Gabriel Crainic

September 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

Service Network Design for Consolidation-based Transportation Advanced Topics[ƚ]

Teodor Gabriel Crainic

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

Abstract. The goal of this and companion chapters is to present a comprehensive overview of the Service Network Design modeling methodology for the planning of consolidation-based freight transportation, focusing on the problem-setting challenges and modeling strategies. This chapter is dedicated to advanced topics, namely, the modeling of demand selection, time, multiple design decisions, resource management, and uncertainty

Keywords: Service Network Design, modeling, freight transportation, consolidation, tactical planning, time dependency, resource management, uncertainty.

Acknowledgements. While working on the project, the author held the UQAM Chair on Intelligent Logistics and Transportation Systems Planning, and was Adjunct Professor, Department of Computer Science and Operations Research, Université de Montréal. He gratefully acknowledges the financial support provided by the Natural Sciences and Engineering Council of Canada (NSERC), through its Discovery grant program, and of the Fonds de recherche du Québec through their infrastructure grants. The author is grateful to Ms. Diane Larin for her invaluable linguistic and editorial assistance, and to Professors Maria Elena Bruni and Pirmin Fontaine, Dr Sara Khodaparasti, and Dr Giacomo Lanza, as well as the Editors and two anonymous referees, for their enriching comments and suggestions.

[ƚ] Part 2 of 2 – Part 1 CIRRELT-2024-30 "Service Network Design for Consolidation-based Transportation – The Fundamentals".

---

* Corresponding author: teodorgabriel.crainic@cirrelt.net

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

1 Introduction

This chapter and its companion (Crainic, 2024b) discuss *Service Network Design* (*SND*) models for the planning of consolidation-based freight transportation. The goal is to present a comprehensive overview of the general SND methodology, focusing on the problem-setting challenges and modeling strategies. The companion chapter focuses on the fundamental SND problem settings and models. This chapter is dedicated to more advanced topics, the modeling of demand selection, time, multiple design decisions and resource management, and uncertainty.

The interested reader may consult a number of general and domain-specific surveys and syntheses (for brevity reasons, only the most recent are mentioned) and the references within: Crainic and Rei (2024); Crainic and Hewitt (2021); Hewitt et al. (2021) (general syntheses), Chouman and Crainic (2021) (railroads), Bakir et al. (2021) (LTL motor carriers), Christiansen et al. (2021) (maritime transportation), and Crainic et al. (2021b) (City Logistics).

The chapter is organized as follows. Section 2 briefly recalls SND problem context addressing consolidation-based freight transportation system and planning. Section 3 states the notation and extends the basic SND formulations to account for demand-selection issues, as well as multiple layers of design. Sections 4, 5, and 6 discuss in some depth the characteristics and modeling of three important extensions of the basic tactical-planning process and SND formulations: time characteristics of demand and services, resource-management concerns, and the explicit consideration of uncertainty, respectively. The chapter concludes in Section 7 with a number of research perspectives we deem important and challenging.

2 SND Problem Context: Consolidation & Planning

Freight transportation constitutes an important economic and social domain, supporting most activities of people and organizations by supplying goods where, when, and how required. It is an equally complex field, transportation activities and flows being the result of the *demand-supply* interactions taking place, within a given political, social, cultural, and economic, environment, among many stakeholders with diverse and often conflicting objectives. To simplify the presentation, and following Bruni et al. (2024), we identify the *shippers*, which generate the multi commodity, origin-destination demand, and the *carriers* supplying the transportation and terminal resources, services, and capacity required to satisfy that demand.

Producers, traders, brokers, importers/exporters, whole- and retail buyers, sellers, and distributors of goods may be described as shippers. The general "carrier" term encompasses the actual modal and intermodal companies physically moving freight (the usual meaning of "carrier"), as well as facility (terminals, warehouses, depots, fulfillment and distribution centers, etc.) operators when independent of the first group. We include logistics service providers (LSPs) in the carrier group. Also known as third/fourth-party logistics firms (3/4LPs, depending on the type of service provided), they generally do not own transportation means, but own or operate facilities. LSPs rather act as intermediaries between shippers and carriers, proposing transportation and logistics services to the former, while contracting with the latter to move the goods received from their customers.

Consolidation is a widely spread strategy aiming to both decrease the costs and increase the efficiency of transportation and logistics, which move a very large and valuable part of the world exchanges, consumer goods in particular, over city, region, country, continent, and inter-continent distances. Consolidation is performed by combining freight of different shippers, with potentially different origins and destinations, and loading them into the same vehicle or container for their complete or partial journeys. It thus reduces the unit shipment cost and the journey time, benefiting all parties involved. Railways, Less-than-Truckload (LTL) motor carriers, shipping companies moving containers on oceans, seas, rivers, and canals, postal services and express couriers, logistics service providers, as well as City Logistics, synchromodal, and Physical Internet systems are prime examples of consolidation-based transportation.

To profitably and efficiently satisfy the requests of many shippers with the same services and resources, carriers organize their operations according to a *transportation plan*, optimizing the resource utilization and deployment through a *service network* that answers the expected/estimated demand. The plan-building process is part of the carrier's mid-term planning and is therefore sometimes called *tactical plan*.

The plan is drawn to be repetitively executed for the duration of a medium-term planning horizon, which we call *season*, for which there is a significant volume of *regular demand*. In terms of volume, regular demand is expected to make up a good part, e.g., 75% - 80%, of the pick demand to be serviced on a "normal" operating day. In terms of consistency, demand, and, hence, service, is expected to be repetitive according to a certain pattern, e.g., every day or week. The plan is thus produced for a given time duration, called *schedule length*, and is to be applied repetitively for the duration of the season. Of course, particular service networks may be built for specific moments, e.g., for week days different from weekends, and be combined to form the repetitive plan and schedule length.

The goal of tactical planning is thus to design service networks and operations that are efficient from the point of view of using the carrier's resources (own, rented, or obtained through outsourcing part of operations), and provide the means to achieve the best trade-off between the shipper requirements and expectations and the carrier pursuit of profitable operations. It involves deciding system-wide the selection and scheduling of services, the consolidation and transfer of freight and vehicles in terminals (as well as on the convoy makeup and dismantling for rail, road, and barge trains), the assignment and management of resources to support the selected services, and the routing of freight of each particular demand through the designed service network. It therefore involves the planning of several different but strongly interrelated hauling and terminal activities, which marshal several types of human and material resources, and are characterized by not-necessarily converging objectives and strong network-wide impacts. Trade-offs among the various components of the transportation system and operations in terms of costs, time, and resource utilization must therefore be achieved as well. SND is the *Operations Research* (*OR*) methodology of choice to address these challenging requirements.

3 Basic SND with Demand Selection & Multiple Layers

Service Network Design, part of the core OR *Network Design* combinatorial-optimization family of problems and methodology, has been largely applied to consolidation-based freight transportation, as witnessed by a rich body of literature (summed up in Crainic et al., 2021a).

Crainic (2024b) presents the modeling challenges, notation, and formulations for the basic SND problem setting, namely, static, deterministic, homogeneous demand type, single design layer, and no consideration for management of resources or revenues. We recall those necessary concepts in this section, by extending the basic problem setting to address what may be called demand-aware SND (Section 3.2) and multi-layer design (Section 3.3). We initiate the section by briefly recalling core elements of SND modeling. Here as in the rest of the chapter, we follow the notation of Crainic and Hewitt (2021) and Crainic (2024b).

3.1 Core SND Components

We discuss the physical system, the demand, and the potential service network of SND models. Resources, and their management-concerned assignment to services, enrich this setting and are discussed in Section 5.

Physical network. Carriers providing consolidation-based services operate on single-mode or multimodal infrastructure networks made up of terminals linked by physical (roads and rail lines) or conceptual (maritime and air corridors) connections. Transfers between modes take place at intermodal terminals. Freight shipments, systems, and networks are identified as intermodal when freight is moved in a loading unit, a container or trailer, most of the time, without touching the freight itself when the loading unit is transferred from one mode to another.

Two types of terminals are identified for tactical planning, *hubs* and *regional*. The latter term groups terminals where most of the demand from the surrounding regions is brought in to be transported by the carrier, and where the freight coming from other regions terminate their trips before being distributed to their final destinations. Hubs are focal points for consolidation-based transportation systems. While also acting as the regional terminals for their hinterlands, their main role is to *consolidate* the flows in and out of their associated regional terminals for efficient long-haul transportation and economies of scale.

The physical infrastructure system is represented by a *physical network* $\mathscr{G}^{\mathrm{PH}} = (\mathscr{N}^{\mathrm{PH}}, \mathscr{A}^{\mathrm{PH}})$, where node $\eta \in \mathscr{N}^{\mathrm{PH}}$ stands for terminal, hub or regional, while arc $a \in \mathscr{A}^{\mathrm{PH}}$ represents the possibility to move directly between the two terminals represented by its defining origin and destination nodes. Several attributes are associated to the nodes and arcs of $\mathscr{G}^{\mathrm{PH}}$, namely, capacities and costs in terms of vehicles, convoys, containers, or freight volumes. Arc length and infrastructure-quality status are also found.

Demand. Recall that carrier tactical planning and the associated SND models target setting up the organization of the carrier's regular operations to service the estimated/forecast regular demand for transportation of its shipper customers. Demand is thus generally defined as the requests for transportation of a set $\mathscr{K}$ of *origin-destination* (OD) commodities, each *commodity* being an aggregation of shippers with similar characteristics in terms of origin, destination, timing, type, product, handling cost, and fare.

The demand characterization of most SND contributions in the literature concerns a single category of customers, which are strongly believed to require service regularly during the coming season. We identify this category as *contract-based*, with associated commodity set $\mathscr{K}^{\mathrm{C}}$. This demand must be satisfied by the designed service network. Consequently, the total associated revenue R^{C} is assumed to be a given constant, to be ignored in planning and the SND minimization of the total operation costs.

To illustrate the capability of the SND methodology to account for demand segmentation and service differentiation, we define a second demand category, identified as *irregular-potential*, with associated commodity set $\mathscr{K}^{\mathrm{I}}$. This category represents the aggregated volume of estimated demand the carrier receives on a regular basis, while the plan is executed, from shippers without formal understandings. This type of demand may be accepted or not. Let $cat(k)$ stand for the category of demand, $cat(k) = \mathrm{C}$ or I when $k \in \mathscr{K}^{\mathrm{C}}$ or $\mathscr{K}^{\mathrm{I}}$, respectively.

The entire demand set is then $\mathscr{K} = \mathscr{K}^{\mathrm{C}} \cup \mathscr{K}^{\mathrm{I}}$, each commodity $k \in \mathscr{K}$ meaning a request to move a quantity of freight vol_k from its origin $O(k)$ to its destination $D(k)$. Each demand is also characterized by its shipper *class* with respect to service requirements, defined mainly is terms of product-type characteristics (packing, handling, and admissible vehicle types) and rapidity of delivery (standard and express, for example). The unit *fare* (tariff) ρ_k the shipper pays to the carrier is associated to this class. (See, e.g., Bilegan et al., 2022; Taherkhani et al., 2022, for the SND modeling of more complex customer and service differentiations, timing chracteristics, and early/late-delivery penalties.)

Service network. Consolidation-based carriers aim to satisfy demand by organizing their operations into so-called *hub-and-spoke service networks*, each *service* being defined by a (modal) physical route between a pair of origin and destination nodes (terminals). Carriers then

- Move low-volume outgoing loads from regional terminals to their designated hubs using *feeder* services;
- Sort (*classify*) the loads brought to hubs, consolidating them into larger shipments to be routed to other hubs by high-frequency, high-capacity *long-haul* services; Loads may go through more than one intermediary hub before reaching their regional terminal destination, being simply transferred from one service to another or undergoing re-classification and re-consolidation;
- Unload from long-haul services and sort the loads arriving at the last hub on their itineraries, and then move them by feeder services to their destination regional terminal.

Notice that, when the level or value of demand justifies it, long-haul services may be run between a hub and a regional terminal or between two regional terminals. Notice also that, more than one service, of possibly different modes, may be operated between terminals.

Let $\mathscr{G} = (\mathscr{N}, \mathscr{A})$ be the potential service network, defined based on the physical nodes of the system $\mathscr{N}^{\mathrm{PH}}$, and the set of potential services Σ, within the context of the carrier resources, operation rules, economics, and service goals.

A service $\sigma \in \Sigma$ follows a path in the physical network from its origin $O(\sigma)$ to its destination $D(\sigma)$. Several other terminals may be located along this path. A *direct* service passes by these terminals without stopping. The service is then represented as a single arc $a \in \mathscr{A}$, and $\mathscr{A} = \Sigma$ when all potential services are single-leg. This is the case of the basic SND formulation (Crainic, 2024b), and is encountered often in the literature, particularly in LTL motor-carrier applications.

A *multi-leg* service halts at intermediary terminals on its route to drop and pick up loads. When convoys are involved (e.g., rail, road, and barge trains), the service may also stop to pick up or drop off individual or groups of vehicles (e.g., car or blocks for railroads and trailers for LTL motor carriers operating multi-trailer road trains). The service route is then described by the sequences of $n(\sigma)$ terminal stops and $n(\sigma) - 1$ service legs connecting them. A single-leg, direct, service has $n(\sigma) = 2$. Let $\mathscr{N}^{\mathrm{PH}}(\sigma) = \{\eta_i(\sigma) \mid i = 1, \ldots, n(\sigma), O(\sigma) = \eta_1, D(\sigma) = \eta_{n(\sigma)}\}$ be the stop sequence of service $\sigma \in \Sigma$. Then, the *service leg* $l_i^\sigma = (\eta_i, \eta_{i+1})$ is defined as the sub-path connecting the consecutive terminals η_i, $\eta_{i+1} \in \mathscr{N}^{\mathrm{PH}}(\sigma)$ of the route of service σ, with $\mathscr{L}(\sigma) = \{l_i^\sigma, i = 1, \ldots, n(\sigma) - 1\}, \sigma \in \Sigma$. Each service leg makes up an arc in G, i.e., $\mathscr{A} = \mathscr{L} = \cup_{\sigma \in \Sigma} \mathscr{L}(\sigma)$, with $l_i^{\sigma(a)}$ standing for the leg of service $\sigma \in \Sigma$ that defines arc $a \in \mathscr{A}$.

Each service $\sigma \in \Sigma$ is characterized by a fixed cost f_σ, incurred when selecting and operating it, a unit freight-transportation cost c_σ, plus global u_σ and, when appropriate, leg-specific $u_{l_i^\sigma}, i = 1, \ldots, n(\sigma) - 1$, capacities representing the total volume of freight the service may load and haul. When commodity characteristics are relevant, appropriate unit transportation costs c_σ^k and capacities u_σ^k ($u_{l_i^\sigma}^k$) are defined for each commodity $k \in \mathscr{K}$. The service cost and capacity figures are inherited by the corresponding arc $a \in \mathscr{A}$, yielding $c_a, c_a^k, k \in \mathscr{K}$, and $u_a, u_a^k, k \in \mathscr{K}$, respectively. We refer the interested reader to Crainic (2024b) for an in-depth discussion of the concepts and modeling of "costs" and "capacities".

3.2 Basic SND with Demand Selection

Through its decision variables, a SND model represents and integrates major sets of tactical-planning decisions, namely, the design of the service network, the selection of demand, when appropriate, and the utilization of that network to service demand. The latter decisions concern, for each individual demand, the *itinerary*, i.e., the sequence of services, terminals, and terminal operations used to move the corresponding flow. Several itineraries may be used simultaneously for a given demand, when its shipment may be split among several service paths between the respective origin and destination terminals. Define

- The *design* decision variables $y_\sigma \in \mathbb{Z}_+$, represented the *frequency* of service $, \sigma \in \Sigma$,

i.e., the number of times the service is operated during the given period, if selected; $y_\sigma \in \{0,1\}$ in the basic SND models;
- The *demand-selection* decision variables $\zeta_k = 1$, if the irregular demand $k \in \mathscr{K}^{\mathrm{I}}$ is selected, 0 otherwise;
- The *utilization*, or *flow*, decision variables $x_a^k \geq 0$, $a \in \mathscr{A}, k \in \mathscr{K}$, prescribing the amount of commodity k that travels on arc a, i.e., on service leg $l_i^{\sigma(a)}$ of service $\sigma(a) \in \Sigma$.

Formally, then, the basic linear-cost, deterministic, split-demand, arc-based SND formulation with demand selection seeks to

$$\max \quad \mathrm{R}^{\mathrm{C}} + \sum_{k \in \mathscr{K}^{\mathrm{I}}} \rho_k \zeta_k - \left(\sum_{\sigma \in \Sigma} f_\sigma y_\sigma + \sum_{k \in \mathscr{K}} \sum_{a \in \mathscr{A}} c_a^k x_a^k \right) \tag{1}$$

$$\text{s.t.} \quad \sum_{a \in \mathscr{A}_\eta^+} x_a^k - \sum_{a \in \mathscr{A}_\eta^-} x_a^k = d_k \qquad \eta \in \mathscr{N}, k \in \mathscr{K}^{\mathrm{C}}, \tag{2}$$

$$\sum_{a \in \mathscr{A}_\eta^+} x_a^k - \sum_{a \in \mathscr{A}_\eta^-} x_a^k = d_k \zeta_k \qquad \eta \in \mathscr{N}, k \in \mathscr{K}^{\mathrm{I}}, \tag{3}$$

$$\sum_{k \in \mathscr{K}} x_a^k \leq u_{l_i^{\sigma(a)}} y_{\sigma(a)}, \qquad a \in \mathscr{A}, \tag{4}$$

$$x_a^k \leq u^k_{l_i^{\sigma(a)}} y_{\sigma(a)}, \qquad k \in \mathscr{K}, a \in \mathscr{A}, \tag{5}$$

$$y_\sigma \in \mathbb{Z}_+ \qquad \sigma \in \Sigma, \tag{6}$$

$$\zeta_k \in \{0,1\}, \qquad k \in \mathscr{K}^{\mathrm{I}}, \tag{7}$$

$$x_a^k \geq 0, \qquad k \in \mathscr{K}, a \in \mathscr{A}, \tag{8}$$

where $\mathscr{A}_\eta^+ = \{(\eta, \eta') \in \mathscr{A}\}$ and $\mathscr{A}_\eta^- = \{(\eta', \eta) \in \mathscr{A}\}$ define the sets of outgoing and incoming arcs for node $\eta \in \mathscr{N}$, respectively, while $d_k = vol_k$ at the demand origin $\eta = O(k)$, $-vol_k$ at the demand destination $\eta = D(k)$, and zero at all other nodes.

The objective function (1) maximizes the net profit computed as the difference betwwen the profit of the contract-based and selected irregular shipper demands, and the total cost of operating the system equal to the sum of the fixed costs of the designed service network and the variable cost of transporting commodities using the selected services. Equations (2) and (3) are often referred to as *flow-balance* constraints and ensure that all of a commodity's contract-based and selected-irregular demand, respectively, departs from its origin, arrives at its destination, and departs any other locations at which it arrives. The expression on the left-hand side of the *linking* constraints (4) computes the total flow traveling on arc $a \in \mathscr{A}$, whereas the expression on the right-hand side gives the global arc capacity provided by the corresponding service leg. The commodity-disaggregated linking constrains are given by (5). Constraints (6) - (8) define the variable domains.

An equivalent path-based formulation is obtained by defining explicitly the set of possible itineraries Π^k of commodity $k \in \mathscr{K}$ on the potential service network. Let $\mathscr{A}_\pi^k$ hold the sequence of arcs $a \in \mathscr{A}$ making up the itinerary $\pi \in \Pi^k$, and let the Kronecker delta coefficients $\delta_a^{\pi k}$ define this path, i.e., $\delta_a^{\pi k} = 1$ when $a \in \mathscr{A}_\pi^k$, 0, otherwise. The unit itinerary flow cost is then defined as $c^k_\pi = \sum_{a \in \mathscr{A}^k_\pi} c^k_a, k \in \mathscr{K}$ (and $x^k_a = \sum_{\pi \in \Pi^k} \delta^{\pi k}_a h^k_\pi, a \in \mathscr{A}, k \in \mathscr{K}$). The path-based SND formulation may then be written as

$$\max \quad \mathrm{R}^{\mathrm{C}} + \sum_{k \in \mathscr{K}^{\mathrm{I}}} \rho_k \zeta_k - \left(\sum_{\sigma \in \Sigma} f_\sigma y_\sigma + \sum_{k \in \mathscr{K}} \sum_{\pi \in \Pi^k} c^k_\pi h^k_\pi \right) \tag{9}$$

$$\text{s.t.} \quad \sum_{\pi \in \Pi^k} h^k_\pi = d_k \qquad k \in \mathscr{K}^{\mathrm{C}}, \tag{10}$$

$$\sum_{\pi \in \Pi^k} h^k_\pi = d_k \zeta_k \qquad k \in \mathscr{K}^{\mathrm{I}}, \tag{11}$$

$$\sum_{k \in \mathscr{K}} \sum_{\pi \in \Pi^k} \delta^{\pi k}_a h^k_\pi \leq u_{l_i^{\sigma(a)}} y_{\sigma(a)}, \qquad a \in \mathscr{A}, \tag{12}$$

$$\sum_{\pi \in \Pi^k} \delta^{\pi k}_a h^k_\pi \leq u^k_{l_i^{\sigma(a)}} y_{\sigma(a)}, \qquad k \in \mathscr{K}, a \in \mathscr{A}, \tag{13}$$

$$y_\sigma \in \mathbb{Z}_+ \qquad \sigma \in \Sigma, \tag{14}$$

$$\zeta_k \in \{0,1\}, \qquad k \in \mathscr{K}^{\mathrm{I}}, \tag{15}$$

$$h^k_\pi \geq 0, \qquad \pi \in \Pi^k, k \in \mathscr{K}, \tag{16}$$

The arc and path-based formulations (1) - (8) and (9) - (15), respectively, address a large gamut of problem settings. First, they generalize the classic basic SND problem settings and models. (Dropping the irregular-de-mand class, eliminates constraints (3) (respectively, (11)) and (7), and turns the total revenue into a constant that may be eliminated from the objective function, which is then to be minimized.) Second, they may support strategic carrier planning in selecting markets and customers. Third, these formulations make up the SND modeling framework when revenue management strategies are implemented and used by carriers in actual operations (Bilegan et al., 2022). Finally, the modeling framework also addresses the cases of logistics service providers, acting as intermediaries between shipper demand and carrier supply, and of multi-stakeholder systems, City Logistics, synchromodal, and Physical Internet, implementing some form of collaborative decision-making when sharing services and resources (Taherkhani et al., 2022).

3.3 Multi-layer SND

It is noteworthy that railroads generally implement a double consolidation policy, which groups (classifies) cars into *blocks*, which are then consolidated to make up *trains*. More precisely, loaded and empty cars, with different origins and destinations, being present simultaneously in the same terminal, are sorted and grouped into a block, which is then moved as *a single unit* by a series of trains until its destination, where it is broken down, the cars being either delivered to their final consignees or classified again into new blocks. More than one reclassification may make up the itinerary of a given demand flow. Blocks, on the other hand, travel on a series of train services, being simply *transferred* from one train to another at intermediary stops on their routes (Chouman and Crainic, 2021).

This strategy requires SND formulations addressing multiple layers of consolidation, which is not the case for the basic models presented above and in Crainic (2024b). Similar modeling requirements arise when resource-management concerns are part of tactical planning (Section 5), as well as in applications to motor-carrier platooning. SND models based on *multi-layer* service networks address these issues.

Multi-layer service networks display the particular characteristic of an arc in a given decision layer being defined with respect to a set of arcs, often making up a path or a cycle, in another decision layer. The rail case illustrates the concept as each potential block in the block layer is defined in terms of the path of arcs in the service layer that would transport it, if selected. Such interwoven definitions imply several connectivity relations and requirements in terms of both design and flow-distribution decisions, yielding rich *Multi-layer Network Design* (*MLND*) formulations raising challenging algorithmic issues (Crainic, 2024a, and references within).

Let $\mathscr{L}$ be the set of layers of multi-layer network $\mathscr{G} = \bigcup_{l \in \mathscr{L}} \{\mathscr{G}_l\}$, where $\mathscr{G}_l = (\mathscr{N}_l, \mathscr{A}_l)$ is the network on layer $l \in \mathscr{L}$, with $\mathscr{N}_l$ and $\mathscr{A}_l$ the corresponding sets of nodes and arcs. Let $l, l' \in \mathscr{L}$ be a couple of *(supporting, supported*) layers of $\mathscr{G}$ (e.g., (train, block)), coupled by an arc definition specifying how an arc in *supported* layer l' is related to a subset of *supporting* arcs in layer l (e.g., the supporting train service arcs form the path defining the supported block). For simplicity of presentation, we assume that all arcs in $\mathscr{G}$ are design arcs, and that a single set of OD demands $\mathscr{K}$ is defined on a given layer. The SND already-defined notation applies, but has to be adjusted with the appropriate layer index, including for the decision variables $y_{al} = 1$ if arc $a \in \mathscr{A}_l$ of layer l is selected, 0, otherwise; and x^k_{al} indicating the quantity of demand $k \in \mathscr{K}$ assigned to arc a of layer l. The objective function and constraints of the previous SND models are the basis of the MLND formulation, by summing the cost term over all layers and instantiating the constraints for the specific layer where the demand is defined.

The MLND model is completed by a set of constraints (17) corresponding to the design (e.g., selection of an arc is), flow, or attribute connectivity requirements proper to the multi-layer network design application at hand. Two examples: 1) Design connectivity constraints (18) state that all the supporting arcs $\alpha \in \mathscr{A}_l(al')$ must be selected in order for a supported arc $a \in \mathscr{A}_{l'}$ to be eligible for selection; 2) Flow connectivity constraints (19) address the case when the demand $\mathscr{K}$ is defined on the supported layer, and the commodity flow on a supporting arc equals the sum of that commodity flows on all its supported arcs (Crainic, 2024a).

$$(\mathbf{y}, \mathbf{x}) \in (\mathscr{Y}, \mathscr{X})_{ll'}, \qquad l, l' \in \mathscr{L}. \tag{17}$$

$$y_{al'} \leq y_{\alpha l}, \qquad \alpha \in \mathscr{A}_l(al'), a \in \mathscr{A}_{l'}, (l, l') \in \mathscr{C}. \tag{18}$$

$$x^k_{\alpha l} = \sum_{a \in \mathscr{A}_{l'}(\alpha l)} x^k_{al'}, \qquad \alpha \in \mathscr{A}_l,\ k \in \mathscr{K}, (l, l') \in \mathscr{C}. \tag{19}$$

4 Modeling Time

Recall that, the length of the tactical planning horizon is determined by the homogeneity of that time duration, in terms of regular demand and service activities, in a stable environment. This length varies with the carrier type, but may also vary with the climate, e.g., between dry and rainy or between cold (winter) and warm (no-winter) seasons. Recall also that, the tactical-plan schedule length is much shorter than the season length, being largely determined by the repetition pattern of the regular demand. The activities planned for this schedule length are then repeatedly executed for the duration of the tactical planning horizon.

The basic SND formulations discussed previously and in Crainic (2024b) address situations where there are no variations in demand for the duration of the schedule and season lengths. Either the schedule length is short and one assumes that everything "happens simultaneously", or the demand arrivals and service departures are assumed to be uniformly distributed over the schedule length for longer time spans. Problem settings and SND models that do not integrate time-related attributes, at least not explicitly, are qualified as *static*. Notice, however, that, the SND models involving service frequencies (Section 3), which are identified as "static", are modeling time and activities in time nonetheless. The implicit modeling of time is performed by assuming that the multiple executions/departures of the service are equally spread out over the schedule length. Frequencies may be modeled as decisions (Section 3) or extracted out of the optimization results (e.g., Powell and Sheffi, 1983).

Tactical planning problem settings and SND formulations are increasingly and consistently addressing the variations of demand in time and the scheduling of the services selected to answer the transportation requests. *Time-dependent* (the term *time-sensitive* is also used) settings and models explicitly identify time-related attributes of demand, system components, and activities for the duration of the schedule length. *Scheduled Service Network Design* (*SSND*) models are time-sensitive SND formulations then target the selection of *scheduled* services to support decisions related to when services, resources, and freight leave and arrive at terminals on their respective routes and itineraries.

The most frequent time-related attributes of demand are the *availability time* $\alpha(k)$ at the origin $O(k)$ and the *due date* $\beta(k)$ at the destination $D(k)$ of each commodity $k \in \mathscr{K}$. The latter is sometimes refined as an interval around an ideal delivery date, the lower and upper limits of the interval modeling the earliest and latest time the consignee is ready to accept delivery; penalty costs may be associated to these limits and late deliveries. The $\beta(k) - \alpha(k)$ amplitude is the major component of the customer class, reflected into the associated fare.

Services are similarly characterized by a *schedule* indicating the departure and arrival times, $\alpha(\eta_i), i = 1, \ldots, n(\sigma) - 1$, and $\beta(\eta_i), i = 2, \ldots, n(\sigma)$, respectively, at each of the terminals $\eta_i \in \mathscr{N}^{\mathrm{PH}}(\sigma)$ on their routes. The corresponding schedule of leg $l_i^\sigma = (\eta_i, \eta_{i+1})$ is thus departure at $\alpha(l_i^\sigma) = \alpha(\eta_i)$ and arrival at $\beta(l_i^\sigma) = \beta(\eta_{i+1})$. Services are further characterized by a total duration $\tau(\sigma)$, that includes the time spent in terminals and the traveling time associated to each leg $\tau(l_i^\sigma)$. Schedules may be strict, as for most European and Canadian railroads, somewhat flexible (e.g., by specifying a day and time interval), or more of an "indicative" nature, the schedule being eventually modified to account for how much freight is already loaded.

How does one model "time" in the context of SND and tactical planning of consolidation-based carriers? In other words, how does one represent *within the schedule length* the time-attributes of the system elements and the associated events, e.g., demand becoming available at origin, the starting and duration of a classification activity in a terminal, and the scheduled departure of a service?

To answer such questions and capture the time-related characteristics of demand and supply, SSND models are generally defined on a *time-space network* $\mathscr{G} = (\mathscr{N}, \mathscr{A})$, which is typically built by extending the network $\mathscr{G}^{\mathrm{PH}}$ along the dimension of time for the duration of the schedule length. The modal service legs provide the (potential) arcs supporting the movements through space and time of the vehicles and convoys of the services considered, while itineraries perform the same role for the transportation of time-dependent demand. Notice that, unless otherwise specifically indicated, most SSND contributions in the literature assume the demand amplitude and the service duration to be contained within the schedule length.

A time-space network is often built by partitioning the schedule length into non-overlapping periods of time. Then, as it is standard modeling practice in multi-period optimization, all activities taking place at a terminal during a period are assumed to occur at the same time instant, be it the beginning (most often) or the end of the period. This approach is also known under the term *time discretization*. The granularity of the discretization and the definition of each period are normally governed by the characteristics of the system studied and operation practice. Most applications in the literature implement the classical approach, first introduced by Ford and Fulkerson (1958), according to which all periods are of the same length and apply to all the nodes of the network. As discussed in Crainic and Hewitt (2021), however, this approach is not always appropriate.

The schedule-length partition does not need to be the same at all terminals, nor, in fact, the same at all times for the same terminal. The relative importance of the terminal with respect to the overall work load generally points to the need to have a fine granularity for high-utilization terminals, hubs in particular. Many services start, stop, or terminate, at such terminals, over most of the schedule length, and many demand flows need to be handled. In contrast, smaller regional terminals may have to operate at less intensive levels and at certain time moments only.

We therefore define the set $\mathscr{T}(\eta) = \{1, 2, \ldots, \iota(\eta)\}$ of periods discretizing the schedule length T for node (terminal) η; $\mathscr{T} = \cup_{\eta \in \mathscr{N}^{\mathrm{PH}}} \mathscr{T}(\eta)$. The discretization is performed by a sequence of time instants $t_i^{\eta}, i = 1, \ldots, \iota(\eta) + 1$, each period $\iota \in \mathscr{T}(\eta)$ being defined by its starting and finishing time instants, t_{ι}^{η} and $t_{\iota+1}^{\eta}$ (i.e., the starting of the next period), respectively. The node set of the time-space network $\mathscr{G}$ may then be defined as $\mathscr{N} = \{(\eta, t_i^{\eta}), \eta \in \mathscr{N}^{\mathrm{PH}}, i = 1, \ldots, \iota(\eta)\}$, including copies of all physical-terminal nodes at all relevant time instants.

Figure 1 illustrates this modeling approach for a tiny network of four terminals and a schedule of length $T = 4$. The same discretization into four equal-length periods, defined by five time instances, applies to all terminals. All events are assumed to take place at the beginning of the respective period. Three services are identified by color and line pattern on the physical network at the left of the figure. The service schedules display the terminal and periods of departure and arrival of each service leg. Short stop times at intermediary terminals are captured by the departure time from the node being equal to the arrival time of the previous leg at the same node, as illustrated by service σ_3 at terminal C. The stop duration is normally included in $\tau(\sigma)$. The corresponding deployment of each service in time and space, according to its schedule, is displayed on the time-space network at the right of the figure.

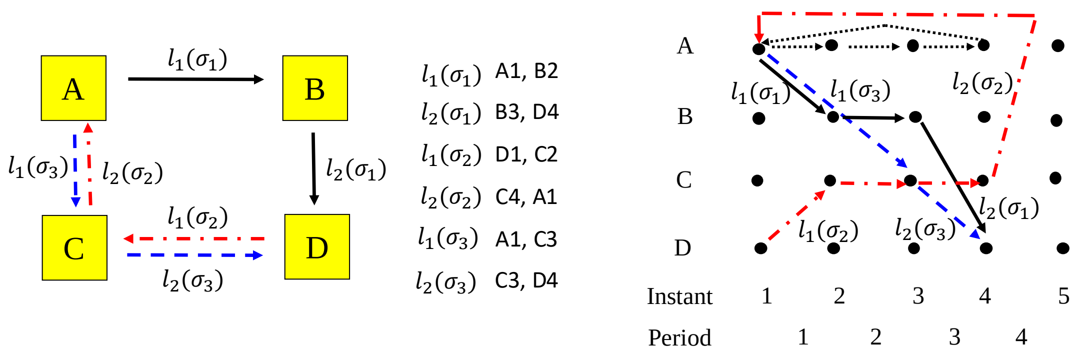

Figure 1: Time-space network representation

The arc set of the time-space network $\mathscr{A}$ consists of *moving* and *holding* arcs. Moving arcs correspond to the scheduled service legs. Specifically, the moving arc standing for leg $l_i^\sigma = (\eta_i, \eta_{i+1}), i = 1, \ldots, n(\sigma), \sigma \in \Sigma$, is defined as $a = ((\eta_i, \alpha(\eta_i)), (\eta_{i+1}, \beta(\eta_{i+1}))$, representing the departure of the service leg from its origin terminal η_i at time instant $\alpha(\eta_i)$ (starting time of a certain period $\iota(l_i^\sigma) \in \mathscr{T}(\eta)$) and arriving at its destination terminal η_{i+1} at time $\beta(\eta_{i+1})$ (starting time of period $\iota(l_i^\sigma) + \tau(l_i^\sigma)$).

Holding arcs represent the possibility for freight or resources to wait, to be "held", at terminals for one period. Such an arc is thus of the form $a = ((\eta, t_i^\eta), (\eta, t_{i+1}^\eta)), \eta \in \mathscr{N}^{\text{PH}}, i = 1, \ldots, \iota(\eta) + 1$. There are no fixed costs associated with holding arcs. A number of capacity and unit cost parameters may be defined, however, to represent the handling or warehousing capabilities of the terminal and costs of this handling or of keeping resources idling.

Figure 1 illustrates six moving arcs for the six service legs, two holding displaying the two-period stop service σ_2 performs at terminal C, and the four holding arcs linking successive representations in time of terminal A (the other holding arcs at terminals B, C, and D are not shown for clarity's sake).

Figure 1 includes two arcs, one moving from terminal C at period 4 (instant 4) to terminal A at period 1 (instant 1), and the holding arcs from time 4 to time 1 at terminal A, which seem to go backward in time. The apparent time travel is only an illusion, however, produced by the graphic display of a modeling device to account for, on the one hand, the repetitiveness of the tactical plan answering the regularity of demand over the season and, on the other hand, the fact that some demands and services may not have both their initial and terminal instants within the schedule length; some may start during the previous application of the plan and terminate currently; others, start during the current application of the plan, but terminate during

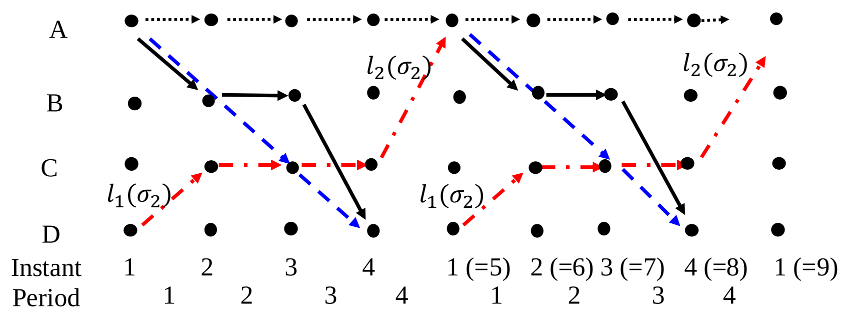

Figure 2: Time-space network representation over two schedule-length duration

the following one. Service σ_2 (red dash-dot line) and the holding arcs at terminal A in Figure 2 illustrate this. Obviously, the unfolding of the time-space network occurs for the complete season duration. These issues are addressed by having the moving and holding arcs modeling activities, which would end during the next application of the plan, *wrap-around*. To avoid the apparent time-travel paradox, the time computations are performed $modulo(T)$ (see, e.g., Crainic and Hewitt, 2021, for details). Note that, the minimization of the transportation costs, together with the commodity due-date constraints, are generally sufficient to avoid cycling in the time-space network.

Figure 3 illustrates the initial and terminal activities of demand itineraries. Notice that, when the availability and due-date times of a commodity are strict, meaning that those operations must be performed at the times given, one could associate the commodity's origin and destination to respective nodes in the time-space network. The modeling approach illustrated in the figure is more general, however. It introduces commodity origin and destination nodes connected to the time-space network through arcs standing for the costs, delays, and penalties (when appropriate) associated to the particular time instances when cargo-handling may start at the origin (left side of the figure) and freight may arrive at destination (right of the figure), respectively. Notice that this approach 1) enforces the timing requirements of each particular commodity, including penalties for early or late handling; 2) facilitates the representation of activities performed at different times on parts of the demand, when demand-splitting is allowed; 3) may be implemented by grouping the commodity nodes within super-source and super-sink nodes, the commodity-specific timings and costs being associated to the connecting arcs.

Mini-time-space networks may be used to appropriately model each terminal in time, these mini-networks being then connected through the services arriving and departing at the node during the interval. Figure 4 illustrates such a node representation. It shows three services arriving at the terminal, two of which are at their destination (inbound full red arrows), the third being at an intermediary stop (dashed red arrow). Two outbound services are displayed, one at its origin (outbound full red arrow), the other departing after its scheduled stop. The figure also displays arcs modeling particular terminal activities including, 1) cargo (and block for a railroad yard) transfer between inbound and outbound services (dotted blue arrows); 2) the possibility for freight on the stopping service to stay on board (the dashed red arrow); 3) cargo at destination being unloaded for delivery (outbound dotted black arrow); 4) the possibility to

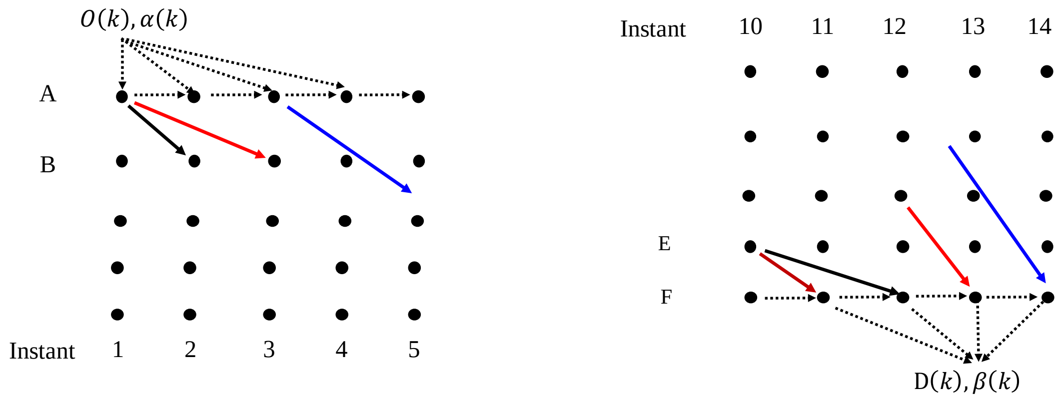

Figure 3: Freight in and out on time-space network representation

unload cargo (or dismantle blocks at destination) for classification and consolidation prior to being put on a departing service (full black arrows); 5) freight at origin becoming available and either going through the classification and consolidation operation, or being put on hold until this operation is to be performed (dotted black arrows). Notice that a continuous-time representation is implied for the activities modeled by the arcs of a mini-network, the particular duration of each being an attribute of the arc. A discrete-time representation proper to the particular terminal and time interval considered is also possible.

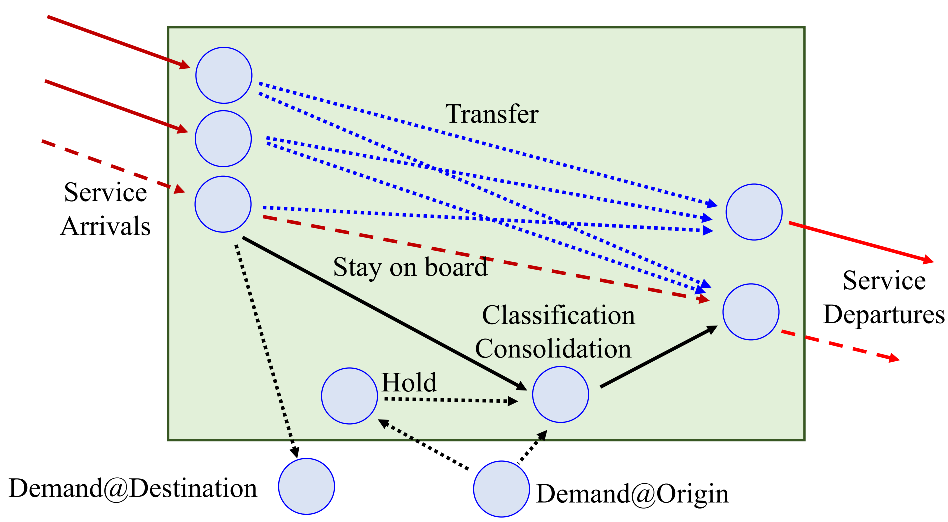

Figure 4: Mini time-space network

Notice that, strict partitioning of the schedule-length is not necessarily required. One may rather use the arrival and departure times of services at terminals to create the time instants of the time-space network, without explicitly defining periods. Then, the physical nodes are duplicated at relevant time instants only, i.e., when the event takes place at the terminal. The availability time of each demand at its origin terminal may be “projected" on the first time instant following arrival. The super-source and super-sink modeling of demand-timing requirements may be used in this setting as well. Figure 5 illustrates this *schedule-based discretization* for two terminals and one service. Three legs are drawn, as well as two stops of not-necessarily equal duration. The arrival and departure schedules make up the time instances of this tiny example. Notice the two holding arcs, dotted lines, parallel to the service stops. They illustrate the unloading and loading of freight, before arriving at destination at the super sink node and after entering the system from the super source node, respectively. One still models on a time-space network in such a representation, but determining an appropriate granularity is not a real issue.

The SSND formulations take the form of the corresponding basic SND models applied to the time-space network $\mathscr{G}$ as defined above. The SSND considers the same sets of decision variables, selecting scheduled services, with frequencies, when several departures may take place within a period, and building itineraries in the service network for demand-flow distribution.

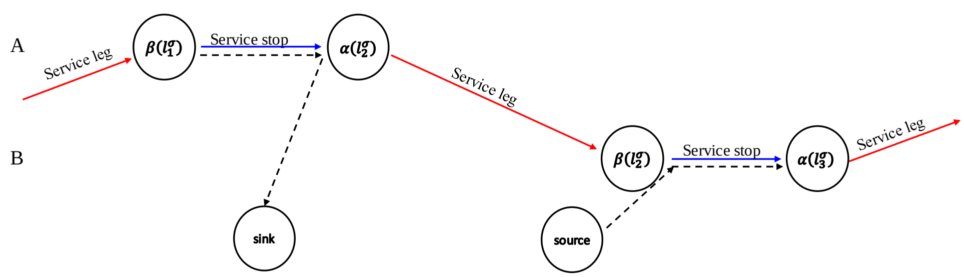

Figure 5: Time-space network with schedule-based discretization

We conclude this section noticing that the representation of time still raises major challenges, notably related to the significant increase in problem size due to discretization. Combining a coarse-granularity-based schedule-length discretization and a frequency-based service selection within the period is an approach to address this issue. Appropriate for short schedule lengths (e.g., a working day as encountered in many LTL applications), it yields periods of generally unequal length, which may be tailored to fit the arrival pattern of demand and the desired pattern of departures to ensure on-time arrival at the next stops of outbound services.

5 Addressing Resource Management Concerns

Material and human resources are needed to operate services and terminals. Notice that the term “asset” is often used in industry and part of the literature. We prefer the term *resource* to identify both the human resources and the material assets of the system.

Two main categories of physical resources are generally considered, *moving* and *terminal*. The former include vehicles, traction/power units, and loading/hauling units (e.g., trailers, railcars, and containers). Terminal-handling resources may be more of less fixed, e.g., quay cranes in port terminals, or mobile, e.g., yard locomotives in railroad networks. *Crews* are required to operate services and terminals. The training and qualification of the crew determines the type of material resources and services the crew can handle.

Assigning human resources to services provides the means to execute them (except when automated vehicles are providing the complete service), while assigning material resources defines their characteristics. Power units determine how much can be hauled in terms of combined load and vehicle weight on each type of physical-network arc, while loading/hauling units determine the capacity of the service. On the other hand, resources are generally costly and in short supply. Consequently, the resource management and resource-to-service assignment issues are very important for the definition, optimization, and performance of the carrier and its service network. In the following, we focus on integrative approaches, aiming at accounting for resource-management issues within SND planning formulations. Crainic and Rei (2024) presents a detailed survey of the evolution of the field.

The initial motivation of research in this field came from the need to reposition resources that completed their assignments and which are needed at terminals different from their current locations. Indeed, trade is unbalanced in the nature, volume, and value of the goods exchanged and, therefore, so are the movements of the resources providing the means for transportation. One thus observes surpluses of certain resources at some terminals and deficits of the same resources at others. Resource *repositioning*, also called *balancing*, is performed to address this issue and make the system ready for the next round of activities.

A first integrative approach focuses on the resource flows at the local level of each terminal by, either including decision variables capturing the balancing flows moving between terminal pairs, or integrating the resource type into the definition of the service it supports. This strategy is generally identified as *design-balanced* SND. In its original form, it addresses problem settings where each service requires a unit of resource (vehicle) only. The idea is to ensure that the number of inbound services (hence, resources) at a node equals the number of outbound services at the same node. The set of node-degree constraints (20) are therefore added to the SND formulations.

$$\sum_{a \in \mathscr{A}_\eta^+} y_a - \sum_{a \in \mathscr{A}_\eta^-} y_a = 0, \ \ \forall \eta \in \mathscr{N}. \tag{20}$$

This approach may be extended in several ways. First, the service-design variables may be multiplied by an appropriate factor when more than one unit of resource is associated to the service, ensuring the balance of the respective resource at the nodes of the network. Second, more than one type of resource may be considered by instantiating a set of constraints similar to (20) for each type. Note that, additional constraints may be required to govern inter-resource relations. Consider, to illustrate, that several container-carrying railcars of different types may be used simultaneously to make up a train. Among other characteristics, the railcar type is determined by the number of platforms on which containers may be loaded and the corresponding car length. Then, one must ensure that the length of the railcar combinations planned to move together on the same service does not exceed the train maximum permitted length.

A second, more comprehensive, integrative approach comes from the observation that models incorporating design-balance type of constraints only do not address many relevant issues.

Thus, e.g., they do not recognize that resources need to periodically return to their specific home-base terminals, Moreover, they are not easily extended to account for other considerations such as the size of the fleet. *Cycle-based* formulations, building on the observation that resources move according to cycles on the (potential) service network, address these issues. Resource cycles are anchored at the home-base terminal to which the resource is assigned. The cycles may start at different periods during the schedule length and be of particular duration, limited in time by the need to return to their home base for inspection and maintenance.

The basic problem setting considers a single resource type. Let $\Theta = \{\theta\}$ be the set of feasible cycles the units of the resource considered may perform, f_θ the "fixed" cost of selecting and operating the resource cycle $\theta \in \Theta$, and δ_θ^σ the cycle-to-service assignment indicator, where $\delta_\theta^\sigma = 1$ if the resource performing cycle $\theta \in \Theta$ may support service $\sigma \in \Sigma$, and 0 otherwise. Define the binary decision variable $y_\theta = 1$, if cycle $\theta \in \Theta$ is selected, and 0 otherwise. The basic SSND with single-resource cycle management then becomes

$$\max \quad \mathrm{R^C} + \sum_{k \in \mathscr{K}^1} \rho_k \zeta_k - \left(\sum_{\sigma \in \Sigma} f_\sigma y_\sigma + \sum_{\theta \in \Theta} f_\theta y_\theta + \sum_{k \in \mathscr{K}} \sum_{a \in \mathscr{A}} c_a^k x_a^k \right) \tag{21}$$

subject to constraints (2) - (8) enriched with

$$y_\sigma \leq \sum_{\theta \in \Theta} \delta_\theta^\sigma y_\theta, \;\; \sigma \in \Sigma, \tag{22}$$

$$y_\theta \in \mathbb{Z}_+, \;\; \theta \in \Theta, \tag{23}$$

where (21) maximizes the net profit accounting for the selection and operation costs of services and resources, as well as for the cost of moving the demand flows. Constraints (22) link the selection of services and the resources required to operate them. The other basic formulations of Sections 3 and 4 may be similarly extended. Notice that, many resource management considerations, such as the cycle duration, are enforced during the *a priori* or dynamic cycle-generation procedures. Limits on the quantity of resources used may be imposed either at the origin nodes of the groups of cycles (which model the base terminals of the respective resources) or at each period.

A more comprehensive problem setting and SSND formulation considers several types of resources, outsourcing servicing certain OD demands, as well as resource acquisition, allocation, and reallocation decisions. The following two-layer SSND with resource acquisition and management, *SSND-RAM*, formulation is based on Crainic and Hewitt (2021).

The SSND notation is adjusted for multiple resources. Let $\mathscr{R}$ stand for the set of available resources, f_η^r the fixed cost of operating a unit of resource of type $r \in \mathscr{R}$ that is assigned to terminal $\eta \in \mathscr{N}^{\mathrm{PH}}$, and I_η^r the quantity of resources of type r initially assigned to terminal η. Let Θ_η^r be the set of potential cycles a resource of type r assigned to terminal η can execute, $\Theta^r = \cup_{\eta \in \mathscr{N}^{\mathrm{PH}}} \Theta_\eta^r$, and $\Theta = \cup_{r \in \mathscr{R}} \Theta^r$. The cycle-to-service assignment indicator δ_θ^σ links services and resources as previously. In all generality, service costs and capacities vary according to the assigned resource, f_σ^r and u_σ^r, $\sigma \in \Sigma$, $r \in \mathscr{R}$, respectively. Notice that a resource-independent fixed service selection cost, f_σ, may still be associated to a service modeling, e.g., the salaries of the officers of a liner ship. Finally, F^r_σ represents the fixed cost of operating service σ with a third party-owned resource of type r (i.e., outsourcing the service).

The service layer of the time-space SSND-RAM network is composed, as previously, of the scheduled (potential) service and holding arcs. The resource layer models the resource cycles, each defined as a sequence of moving and holding arcs in the service layer. Modeling the resource acquisition and allocation adds two types of nodes and arcs to the resource layer, gathered in sets $\bar{\mathcal{N}}$ and $\bar{\mathcal{A}}$, respectively, all nodes in $\bar{\mathcal{N}}$ being symbolically defined at period 0, before the first period of the schedule length:

Acquisition and assignment. A unique node, A, represents the acquisition of new resources, arcs $(A, \eta_1) \in \bar{\mathcal{A}}$ standing for their allocation to terminal $\eta \in \mathcal{N}^{\mathrm{PH}}$ at the first period of activity at that terminal (assumed to the first one, for simplicity of presentation).

Reallocation. A node η_0 is added for each terminal $\eta \in \mathcal{N}^{\mathrm{PH}}$, the arcs $(\eta_0, \eta'_1) \in \bar{\mathcal{A}}, \eta' \in \mathcal{N}^{\mathrm{PH}}$, connect that node to each terminal at the first period of activity and model the reallocation of resources initially at terminal η to terminal η'.

Let h^r_η be the cost of acquiring a new unit of resource $r \in \mathscr{R}$ and allocating it to terminal $\eta \in \mathcal{N}^{\mathrm{PH}}$ (on arcs (A, η_1)). Let $h^r_{\eta_0\eta'_1}$ be the unit reallocation cost for resource $r \in \mathscr{R}$ from terminal $\eta \in \mathcal{N}^{\mathrm{PH}}$ to terminal $\eta' \in \mathcal{N}^{\mathrm{PH}}$ (on arc (η_0, η'_1), $h^r_{\eta_0\eta'_1} = 0$ for $\eta = \eta'$). The cycle definition is extended over the additional nodes and arcs to capture the acquisition and reallocation activities within the resource-routing decisions. Cycles thus start at nodes in $\bar{\mathcal{N}}$, yielding the set $\Theta^r_{\eta_0}$ of potential cycles a resource of type r can execute out of each terminal $\eta \in \mathcal{N}^{\mathrm{PH}}$.

To streamline the presentation, we formulate the SSND-RAM (24) - (34) for the single-leg-service and no-demand-selection case. The SSND decision variables, $y_\sigma, \sigma \in \Sigma$, and $x^k_a \geq 0, a \in \mathscr{A}, k \in \mathscr{K}$, are also defined for the SSND-RAM, with the following additional decision variables:

- $y^r_\sigma = 1$, if service $\sigma \in \Sigma$ is outsourced to a third party-owned resource $r \in \mathscr{R}$, 0, otherwise;
- $z^r_\theta = 1$, if cycle $\theta \in \Theta^r_\eta, r \in \mathscr{R}$, is selected, 0, otherwise;
- w^r_η, number of new units of resource $r \in \mathscr{R}$ acquired and assigned to terminal $\eta \in \mathcal{N}^{\mathrm{PH}}$;
- $w^r_{\eta_0\eta'_1}$, number of units of resource $r \in \mathscr{R}$ reallocated between terminals η and η' in $\mathcal{N}^{\mathrm{PH}}$.

$$\min \sum_{r\in\mathscr{R}} \sum_{\eta\in\mathcal{N}^{\mathrm{PH}}} h^r_\eta w^r_\eta + \sum_{r\in\mathscr{R}} \sum_{(\eta_0,\eta'_1)\in\bar{\mathcal{A}}} h^r_{\eta_0\eta'_1} w^r_{\eta_0\eta'_1} + \sum_{r\in\mathscr{R}} \sum_{\eta\in\mathcal{N}^{\mathrm{PH}}} f^r_\eta \sum_{\theta\in\Theta^r_\eta} z^r_\theta \tag{24}$$

$$+ \sum_{\sigma\in\Sigma} \left(f_\sigma y_\sigma + \sum_{r\in\mathscr{R}} f^r_\sigma \sum_{\theta\in\Theta^r} \delta^\sigma_\theta z^r_\theta \right) + \sum_{\sigma\in\Sigma} \sum_{r\in\mathscr{R}} F^r_\sigma y^r_\sigma + \sum_{k\in\mathscr{K}} \sum_{a\in\mathscr{A}} c^k_a x^k_a$$

$$\text{s.t.} \quad \sum_{(\eta_0,\eta'_1)\in\bar{\mathcal{A}}} w^r_{\eta_0\eta'_1} = I^r_\eta, \qquad r \in \mathscr{R}, \eta \in \mathcal{N}^{\mathrm{PH}}, \tag{25}$$

$$\sum_{\theta\in\Theta^r_\eta} z^r_\theta \leq w^r_\eta + \sum_{(\eta_0,\eta'_1)\in\bar{\mathcal{A}}} w^r_{\eta_0\eta'_1}, \qquad r \in \mathscr{R}, \eta_0 \in \bar{\mathcal{N}}, \tag{26}$$

$$\sum_{a \in \mathscr{A}_\eta^+} x_a^k - \sum_{a \in \mathscr{A}_\eta^-} x_a^k = d_k, \qquad \eta \in \mathscr{N}, k \in \mathscr{K}, \tag{27}$$

$$\sum_{k \in \mathscr{K}} x_a^k \leq \sum_{r \in \mathscr{R}} u_\sigma^r \left(\sum_{\theta \in \Theta^r} \delta_\theta^\sigma z_\theta^r + y_\sigma^r \right), \qquad a \in \mathscr{A}, \tag{28}$$

$$y_\sigma \leq \sum_{r \in \mathscr{R}} \sum_{\theta \in \Theta^r} \delta_\theta^\sigma z_\theta^r, \qquad \sigma \in \Sigma, \tag{29}$$

$$y_\sigma + y_\sigma^r \leq 1, \qquad \sigma \in \Sigma, \tag{30}$$

$$w_\eta^r, w_{\eta_0 \eta_1'}^r \in \mathbb{Z}^+, \qquad r \in \mathscr{R}, \eta_0 \in \bar{\mathscr{N}}, \eta \in \mathscr{N}^{\text{PH}}, \tag{31}$$

$$z_\theta^r \in \{0,1\}, \qquad r \in \mathscr{R}, \theta \in \Theta^r, \tag{32}$$

$$y_\sigma^r \in \{0,1\}, \qquad r \in \mathscr{R}, \sigma \in \Sigma, \tag{33}$$

$$x_a^k \geq 0, \qquad a \in \mathscr{A}, k \in \mathscr{K}. \tag{34}$$

The objective function minimizes the total cost of the system as the sum of the costs to 1) acquire and allocate new resources; 2) reallocate existing ones; 3) operate resources; 4) select and operate services with owned resources; 5) secure third-party resources; 6) transport the demand freight. Constraints (25) ensure that all resources of type r that are initially allocated to terminal i are either left at i or reallocated. Constraints (26) link the strategic resource acquisition and allocation/reallocation decisions that determine the number of resources available at each terminal with the tactical decision of how many resources from that terminal are to be used to execute services. Constraints (27) and (28) enforce classical network design relations. The former are commodity-specific flow conservation constraints. The latter link the existence of flow on arcs to the selection of services supported by owned or outsourced resources. Constraints (29) indicate that at most one owned resource is used for each owned service, while constraints (30) specify that each service cannot be selected more than once, either supported by the carrier's resources or outsourced. Finally, constraints (31) - (34), define the domains of the variables in the formulation.

6 Addressing Uncertainty

Most planning and network design contributions in the literature assume known values for the demand and supply system parameters, which are not supposed to change for the planning-horizon duration. This translates into fixed figures, which may come from historical and field-knowledge data (one finds a lot of average measures) or single-point estimations, when forecasting methods are used. The problem settings and formulations discussed so far belong to this *deterministic* class. The future is uncertain, however, and the explicit representation of the uncertainty of various system parameters, together with the corresponding plan-adjustment actions and costs when new information becomes available, is increasingly part of the tactical-planning and SND literature.

Accounting explicitly for uncertainty in SND and SSND models aims to address these issues. Discussing uncertainty and network design in any depth is beyond the scope of this chapter, however. The interested reader should refer to the documents mentioned in the Introduction, and references within. In the following, we focus on a number of fundamental stochastic-programming modeling concepts as applied to SND.

Uncertainty is generally classified into one of three types based upon their likelihood and impact (Klibi et al., 2010). *Randomness* refers to events whose likelihood can be described and is reasonably high, but whose impacts can usually be mitigated within normal operations. OD demand-volume fluctuations make up a well-studied example of such uncertainty. The second type, *hazards*, refers to events whose likelihood can be described, but are quite rare, e.g., infrastructure or vehicle failure. The third type, *deep uncertainty*, refers to events whose likelihood can not be described, but may have a significant impact, e.g., a maritime port closing down due to a threat of terrorist attack. Most research on SND has focused on the first type of uncertainty. This uncertainty is modeled by extending one of these deterministic models to a stochastic-programming formulation. The section reflects this state-of-the-art.

The challenge when planning within an uncertain context raises from the combination of "when" decisions are taken and "what is known" at decision time. Tactical planing is performed some time before the season starts and, thus, at a time when information regarding the SND parameter values is incomplete, approximations/estimations being available only. New information becomes available as the plan is repeatedly executed during the season. More precisely, each time the plan is to be executed, the actual values of those parameters are *revealed* (*observed*) and become known. One may ignore this information update and execute the plan, or one may adjust it to better fit the observed situation. Financial and service-quality impacts are expected in all cases. The goal therefore is to build SND models that account, at the planning epoch, for these expected variations and plan adjustments, and mitigate their impacts on the system performance.

Stochastic SND models with recourse are the preferred stochastic-programming paradigm for addressing uncertainty in consolidation-based freight transportation planning. These models are designed to reflect the decision-making dynamics of the planning, execution, and *information-revelation* processes. The models explicitly differentiate between decisions made under uncertainty, when the plan is built, and those made after the new information is revealed, when the plan is to be revised and then executed. Then, the objective of a stochastic SND model is to devise a plan that "optimizes" the unchanging part of the plan together with the expected cost over the planning horizon of repeatedly adjusting and executing it.

This setting inherently leads to modeling multiple *stages*, that is, specific moments in time when information becomes available and decisions are taken. In a general sense, tactical decisions are made during the *first stage* under complete uncertainty. They are thus often referred to as *a priori* decisions. Plan-adjustment operational decisions are then executed in subsequent stages, as the stochastic parameters become progressively observed. They are thus known as *recourse* decisions.

Notice that, increasing the number of stages provides the means for more refined representations of the dynamics of operational decision-making. Yet, they also significantly increase the complexity of the optimization model and the computational challenges of addressing it, particularly when discrete decisions are part of some of the later stages. Determining the number of stages is thus an important modeling and decision-support choice. One should remember, thought, that tactical planning does not require detailed operational representations. One should also observe that not all later-stage decisions, which are part of the output of the tactical model, are actually implemented. The demand itineraries are a classic illustration of this point. Demand flows are part of any design model and are thus optimized when the SND formulation is solved. But, in most cases, demand itineraries are optimized anew when the plan is adjusted to revealed information, with or without the guidance of those determined when solving the tactical SND. Thus, the representation of the operational decision-making process within the tactical SND aims only at approximating the impact of first-stage decisions on the performance of the system over the planning horizon. Most stochastic SND contributions in the literature propose two-stage formulations.

How each stage is defined should reflect the specific planning problem under consideration. Typically, the first stage involves design decisions that specify the services to be operated, as well as the demand-selection and resource-management decisions when relevant, thereby establishing the carrier service network. A broad range of potential recourse actions can be established for the second stage, representing varying degrees of flexibility in adjusting the tactical plan. On one end of the spectrum is the *simple recourse* option, which follows an *observe-and-pay* strategy and involves imposing a penalty proportional to the extent of the plan's infeasibility.

More complex recourse actions are formulated at the network level. A notable strategy within this context is the *network strategy*, which treats the optimization of demand itineraries as recourse actions, while design decisions are taken at the first stage. To illustrate for the (1) - (8) basic SND with demand-selection model, let $(\Omega, \mathscr{F}, \mathbb{P})$ be the probability space associated with a random experiment that reveals the demand volumes $d_k, k \in \mathscr{K}$. The set Ω contains the possible outcomes $\omega \in \Omega$ of the random experiment, $\mathscr{F}$ defines the set of events, and $\mathbb{P}$ is the measure assigning probabilities to the possible outcomes of the random experiment. Let y_σ and ζ_k be the first-stage design decisions, and $x_a^k(\omega), a \in \mathscr{A}, k \in \mathscr{K}, \omega \in \Omega$, be the flow variables of the network recourse. The two-stage stochastic-programming formulation then becomes

$$\max \quad \mathrm{R}^{\mathrm{C}} + \sum_{k \in \mathscr{K}^{\mathrm{I}}} \rho_k \zeta_k - \sum_{\sigma \in \Sigma} f_\sigma y_\sigma - \mathbb{E}_\xi \Big[Q\big(y, \zeta, \xi(\omega)\big) \Big] \tag{35}$$

$$\text{s.t.} \quad y_\sigma \in \mathbb{Z}_+, \sigma \in \Sigma, \tag{36}$$

$$\zeta_k \in \{0,1\}, k \in \mathscr{K}^{\mathrm{I}}, \tag{37}$$

where

$$Q\big(y, \zeta, \xi(\omega) = \min \sum_{k \in \mathscr{K}} \sum_{a \in \mathscr{A}} c_a^k x_a^k(\omega) \tag{38}$$

$$\text{s.t.} \sum_{a \in \mathscr{A}_\eta^+} x_a^k(\omega) - \sum_{a \in \mathscr{A}_\eta^-} x_a^k(\omega) = d_k(\omega) \qquad \eta \in \mathscr{N}, k \in \mathscr{K}^{\mathrm{C}}, \tag{39}$$

$$\sum_{a \in \mathscr{A}_\eta^+} x_a^k(\omega) - \sum_{a \in \mathscr{A}_\eta^-} x_a^k(\omega) = d_k(\omega) \zeta_k \qquad \eta \in \mathscr{N}, k \in \mathscr{K}^{\mathrm{I}}, \tag{40}$$

$$\sum_{k \in \mathscr{K}} x_a^k(\omega) \leq u_{l_i^{\sigma(a)}} y_{\sigma(a)}, \qquad a \in \mathscr{A}, \tag{41}$$

$$x_a^k(\omega) \leq u^k_{l_i^{\sigma(a)}} y_{\sigma(a)}, \qquad k \in \mathscr{K}, a \in \mathscr{A}, \tag{42}$$

$$x_a^k(\omega) \geq 0, \qquad k \in \mathscr{K}, a \in \mathscr{A},. \tag{43}$$

Model (35) - (37) addresses the first-stage design decisions for the service network under complete uncertainty. The objective function (35) combines the deterministic components of the total revenues and the fixed costs for the selected services, and the recourse cost function representing an aggregated measure (expectation) of the anticipated future costs of adjusting / adapting the tactical plan to revealed information. The recourse cost function $Q\big(y, \zeta, \xi(\omega)\big)$, defined for the realization $\omega \in \Omega$, evaluates the total cost incurred to optimally move the realized demand $\xi(\omega) = vol_k(\omega)$ given the selected demand ζ and service network y defined in the first stage.

Recourse actions involving design decisions may also be addressed. Consider, to illustrate that one could very well leave for the operational phase the selection of irregular demands. This may be modeled by pushing the ζ decision variables and the associate revenues to the second stage. Modifications to the service network selected in the first stage may also be contemplated when new information becomes available. One may slightly shift the schedule within a allowed interval. One may also activate or cancel services when observed demand is much higher or lower that the predicted level, respectively. Such problem settings yield integer-recourse formulations, network design models actually, with significant impact on resource management and freight flows. Research on such topics is still in its infancy.

7 Conclusions and Perspectives

This chapter and its companion (Crainic, 2024b) present a synthesis of the main classes of Service Network Design models aimed at supporting decision-making in planning the activities and managing the resources of consolidation-based freight carriers and systems. The focus is on modeling issues, concepts, and structures of general interest and relevance. We complete the presentation with a number of challenging research areas.

Resource management concerns and SND. Several issues may be identified. 1) We must enhance the modeling of work rules for moving and terminal crews, as well as of maintenance requirements for various types of equipment. 2) More work is required to model services requiring resources of several types, the combinations and potential substitutions being governed by particular compatibility rules. Each type of resource operates according to its own particular management rules (item 1 above), their representation within SND models yielding multi-layer networks with complex connectivity relations. Very little is currently known on how to efficiently address these issues. 3) Resource capacities to load freight directly determine the service capacity. While the classic capacity constraints, which treat volumes as continuously divisible entities, are sufficient in many cases, they may significantly overestimate what a vehicle or convoy may haul when the loading and packing of demand items is important. Very little planning and SND research has yet addressed this issue, mostly in the "simple" single-dimension case and without proper management of the corresponding resources. More work is needed on integrating these aspects for SND models, first in the simple case and, then, for more complex packing situations.

Modeling time. Explicitly addressing time and delay-related issues enlarges and refines the scope of SND models while raising significant modeling and algorithmic challenges. 1) A major research area is the proper definition of time as integrated into SSND formulations. The granularity issue is still not comprehensively addressed. On the one side, there is the issue of adequately identifying the partitions of the schedule length at various terminals and time intervals, and the seamless integration of such partitions into a unified time-space network. On the other hand, the implicit, schedule-based discretization approach has been little studied so far. Moreover, at our best knowledge, no formal comparison has been conducted between the explicit and implicit discretization approaches. 3) Changing the modeling paradigm and considering time implicitly in SND models is a long-time goal in the research community. A few recent contributions (e.g., He et al., 2023; Lanza et al., 2024) open interesting perspectives that invite for a sustained research effort in the area. 4) More studies are needed with respect to the duration of activities and associated delay-related issues, particularly in terminals and when several activities compete for the same restricted capacity. More refined modeling of mini-terminals and the activities herein, their inclusion into time-space networks, and the approximation of delays with linear or non-linear, ideally convex, functions, are three main issues in this area. Together with the issues related to modeling environmental impacts, this points to the need for more research into network design formulations with non-linear objective functions. 5) Finally, we need to generalize these studies to the multi-resource/layer problem settings.

Uncertainty. One may state that research in uncertainty and SND is still in its infancy. 1) Studies are still required to adequately represent demand uncertainty in various problem settings. Furthermore, almost totally overlooked, although of great operational and economic importance, is the uncertainty in travel and terminal-activity times. The solution often adopted in practice of adding large buffers to the planned delivery times is not only scientifically unsatisfactory, but also less and less economically viable and impracticable in many cases (City Logistics and Synchromodality, to name but two examples). Moreover, one should not overlook that both demand and time uncertainty (and heavy correlations) characterize operations, and that their simultaneous presence and interactions should be reflected in the planning models proposed. 2) Most contributions in the literature address business-as-usual cases, when uncertainty can be somewhat easily represented with probability distributions. Other sources of uncertainty exist, however, and should be studied. Reliability and robustness are two such issues, as is resilience, i.e., the capability to rebound following an incident, and the operation plans to perform the recovery and return to a desired state of system and operation behavior.

Advancing in this direction would also lead to a broader exploration of information-revelation mechanisms and multi-stage formulations.

Extending the field of interest. We limit our discussion to two main areas. 1) *Multi-stakeholder* problem settings typically include several organizations, not necessarily of same type (e.g., private carriers and public transit authorities moving freight on some of their tramway and bus vehicles), which engage in some form of cooperation to share resources and coordinate decisions and operations. The accurate definition and modeling within SND formulations of various types of stakeholder interactions, associated cooperation mechanisms, and work, risk, and cost/benefit sharing make up a major research challenge. 2) Most contributions to the literature assume known customer, that is, demand behavior, with respect to economic and service-level criteria. This is true even when uncertainty in these elements is explicitly represented. Or, customers do react to tariffs and require quality-of-service levels and, consequently, so is the demand the carrier will ultimately service and the revenues it can potentially earn. Extending the SND to address such issues requires considering not only a profit-maximizing objective, but also modeling in mathematical terms the behavioral relations between tariffs, service-quality levels, and the willingness of customers to give a carrier their business. Integrating *revenue management* strategies in carrier operations, planning, and associated SND decision-support methodology makes up a major research area. The formulation presented in Section 3 is a small first step in this direction, but more research is needed as one cannot simply transpose people-transport-related results to the freight transport environment (Bilegan et al., 2022). Integrating carrier SND models into larger bi-level formulations studying, e.g., cooperation mechanisms or the impact of governmental regulations, makes up another major research area which has received very little attention so far.

Transport technology is evolving, concerning, in particular autonomous vehicles, platooning, non-carbon-based motorization, alternate and combined air and land-based modes, etc. This evolution is accompanied by changes in operation requirements and strategies. Research is needed to adequately reflect those into SND planning models and methods.

Extending the concern types. 1) The evolution of the planet climate and related people concerns raise major research challenges for OR in general and SND in particular. Consider, for example, the impact of climate change on the infrastructure. Changing the plan according to the season (e.g., the total permuted weight of loaded motor vehicles is lowered in spring, during the thaw period, to preserve the road infrastructure) used to be sufficient in most cases. This is less and less the case. Two examples. First, the permafrost soil is unfreezing putting most infrastructure at high latitudes at very high risks. Second, the large and unusual variations in rainfalls change quite drastically the water levels on rivers and canals jeopardizing normal navigation conditions. From an OR perspective, research is needed to model such phenomena and their impacts and integrate the results into more comprehensive planning formulations. 2) Related to the previous item, but a major research area on its own, is the representation of environmental concerns into the objective functions of our models. On the one hand, the physics of those phenomena are complex and appropriate synthetic representations are needed for tactical-level planning methods. On the other hand, the social impact of transportation on communities should be better reflected in planning and models. The “nuisance” criterion introduced in City Logistics (Crainic et al., 2009) is a first, modest step in this direction.

Solution methods. Addressing large SND models makes up an extremely important and challenging research area. Exploring this area is beyond the scope of the chapters and the book. Hence, we only mention a few particularly interesting avenues: 1) Dynamic generation of services (paths), resource work assignments (cycles), blocks (paths), demand-flow itineraries (paths), and granularity of time-space network; 2) Development of heuristic-type solution methods, matheuristics, in particular, combining exact and meta-heuristic principles, ideally coupled with decomposition and parallel optimization strategies; 3) Development of efficient solution methods for stochastic SND, even for the two-stage formulations of business-as-usual demand uncertainty case, which has been studied the most. Efficient decomposition and scenario-reduction methods (e.g., Rahmaniani et al., 2017; Hewitt et al., 2021, 2022) offer interesting starting points for what should be a significant research effort.

Acknowledgements

While working on the project, the author held the UQAM Chair on Intelligent Logistics and Transportation Systems Planning, and was Adjunct Professor, Department of Computer Science and Operations Research, Université de Montréal. He gratefully acknowledges the financial support provided by the Natural Sciences and Engineering Council of Canada, through its Discovery grant program, and of the Fonds de recherche du Québec through their infrastructure grants. The author is grateful to Ms. Diane Larin for her invaluable linguistic and editorial assistance, and to Professors Maria Elena Bruni and Pirmin Fontaine, Dr Sara Khodaparasti, and Dr Giacomo Lanza, as well as the Editors and two anonymous referees, for their enriching comments and suggestions.

References

I. Bakir, A. Erera, and M.W.F. Savelsbergh. Motor Carrier Service Network Design. In T.G. Crainic, M. Gendreau, and B. Gendron, editors, *Network Design with Applications in Transportation and Logistics*, chapter 16, pages 427–467. Springer, Boston, 2021.

I.C. Bilegan, T.G. Crainic, and Y. Wang. Scheduled Service Network Design with Revenue Management Considerations and an Intermodal Barge Transportation Illustration. *European Journal of Operational Research*, 300(1):164–177, 2022.

M.E. Bruni, T.G. Crainic, and G. Perboli. Bin Packing Methodologies for Capacity Planning in Freight Transportation and Logistics. In T.G. Crainic, M. Gendreau, and A. Frangioni, editors, *Contributions to Combinatorial Optimization and Applications – A Tribute to Bernard Gendron*, chapter 6, pages 115–147. Springer, 2024.

M. Chouman and T.G. Crainic. Freight Railroad Service Network Design. In T.G. Crainic, M. Gendreau, and B. Gendron, editors, *Network Design with Applications in Transportation and Logistics*, chapter 13, pages 383–426. Springer, Boston, 2021.

M. Christiansen, E. Helsen, D. Pisinger, D. Sacramento, and C. Vilhelmsen. Liner Shipping Network Design. In T.G. Crainic, M. Gendreau, and B. Gendron, editors, *Network Design with Applications in Transportation and Logistics*, chapter 15, pages 469–505. Springer, Boston, 2021.

T.G. Crainic. Multi-layer Network Design for Consolidation-based Transportation Planning. In T.G. Crainic, M. Gendreau, and A. Frangioni, editors, *Contributions to Combinatorial Optimization and Applications – A Tribute to Bernard Gendron*, chapter 9, pages 179–205. Springer, 2024a.

T.G. Crainic. Service Network Design for Consolidation-based Transportation – The Fundamentals. In S. Parragh and T. Van Woensel, editors, *Research Handbook on Transport Modeling*, chapter 3. Edgar Elgar Publishing, 2024b. Publication CIRRELT-2024-30.

T.G. Crainic and M. Hewitt. Service Network Design. In T.G. Crainic, M. Gendreau, and B. Gendron, editors, *Network Design with Applications in Transportation and Logistics*, chapter 12, pages 347–382. Springer, Boston, 2021.

T.G. Crainic and W. Rei. 50 Years of Operations Research for Planning Consolidation-based Freight Transportation. Publication CIRRELT-2024, Centre interuniversitaire de recherche sur les réseaux d'entreprise, la logistique et le transport, Université de Montréal, Montreal, 2024.

T.G. Crainic, N. Ricciardi, and G. Storchi. Models for Evaluating and Planning City Logistics Transportation Systems. *Transportation Science*, 43(4):432–454, 2009.

T.G. Crainic, M. Gendreau, and B. Gendron, editors. *Network Design with Applications in Transportation and Logistics*. Springer, Boston, 2021a.

T.G. Crainic, G. Perboli, and N. Ricciardi. City Logistics. In T.G. Crainic, M. Gendreau, and B. Gendron, editors, *Network Design with Applications in Transportation and Logistics*, chapter 16, pages 507–537. Springer, Boston, 2021b.

L.R. Ford and D.R. Fulkerson. Constructing Maximal Dynamic Flows from Static Flows. *Operations Research*, 6(3):419–433, 1958.

Y. He, M. Hewitt, F. Lehuédé, N. Medina, and O. Péton. A Continuous-time Service Network Design and Vehicle Routing Problem. *Networks*, 83(2):300–323, 2023.

M. Hewitt, W. Rei, and S.W. Wallace. Stochastic Network Design. In T.G. Crainic, M. Gendreau, and B. Gendron, editors, *Network Design with Applications in Transportation and Logistics*, chapter 10, pages 283–315. Springer, Boston, 2021.

M. Hewitt, J. Ortmann, and W. Rei. Decision-based Scenario Clustering for Decision-making under Uncertainty. *Annals of Operations Research*, 315:747–771, 2022.

W. Klibi, A. Martel, and A. Guitouni. The Design of Robust Value-Creating Supply Chain Networks: A Critical Review. *European Journal of Operational Research*, 203(2):283–293, 2010.

G. Lanza, T.G. Crainic, M. Passacantando, and M.G. Scutellá. A Continuous Time Physical Graph Based Formulation to Scheduled Service Network Design. In *Optimization in Green Sustainability and Ecological Transition;*, volume 2 of *AIRO Springer Series*, pages 265–275. Springer, 2024.

W.B. Powell and Y. Sheffi. The Load Planning Problem of Motor Carriers: Problem Description and a Proposed Solution Approach with Predetermined Step Sizes. *Transportation Researc Part A: General*, 17(6):471–480, 1983.

R. Rahmaniani, T.G. Crainic, M. Gendreau, and W. Rei. The Benders Decomposition Algorithm: A Literature Review. *European Journal of Operational Research*, 259(3):801–817, 2017.

G. Taherkhani, I.C. Bilegan, T.G. Crainic, M. Gendreau, and W. Rei. Tactical Capacity Planning in an Integrated Multi-stakeholder Freight Transportation System. *OMEGA*, 110, 2022. Article 102628.