



# CIRRELT

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

Interuniversity Research Centre  
on Enterprise Networks, Logistics and Transportation

---

## A Survey and Taxonomy of Multilayer Network Design

Mohammad Rahim Akhavan Kazemzadeh  
Teodor Gabriel Crainic  
Bernard Gendron

March 2019

CIRRELT-2019-11

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

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

[www.cirrelt.ca](http://www.cirrelt.ca)

# A Survey and Taxonomy of Multilayer Network Design

Mohammad Rahim Akhavan Kazemzadeh<sup>1,2,\*</sup>, Teodor Gabriel Crainic<sup>1,3</sup>,  
Bernard Gendron<sup>1,2</sup>

<sup>1</sup> Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT)

<sup>2</sup> Department of Computer Science and Operations Research, Université de Montréal, P.O. Box 6128, Station Centre-Ville, Montréal, Canada H3C 3J7

<sup>3</sup> Department of Management and Technology, Université du Québec à Montréal, P.O. Box 8888, Station Centre-Ville, Montréal, Canada H3C 3P8

**Abstract.** In multilayer network design, decisions are represented by different potential networks, each at a given layer. In each layer, flow requirements for a set of commodities must be satisfied. To route the commodities, appropriate arcs have to be selected (or opened) in each layer. There are several types of coupling constraints between the layers. For example, to open an arc in a particular layer, supporting arcs in another layer have to be opened. Applications of the multilayer network design problem can be found in the fields of transportation and telecommunications. Although this is an important class of problems in combinatorial optimization, to the best of our knowledge, there is no survey on the topic which covers extensively multilayer network design problems. In this paper, we propose the first classification and a state-of-the-art survey of multilayer network design problems. The survey focuses on applications in transportation and telecommunications, as well as on solution methods. We also propose a general modeling framework that encompasses most multilayer network design problems found in the literature.

**Keywords.** Networks, multilayer network design, transportation, telecommunications, combinatorial optimization.

**Acknowledgements.** While working on this project, the second author was Adjunct Professor with the Département d'informatique et de recherche opérationnelle (DIRO), Université de Montréal. The first author gratefully acknowledges the support of the Université de Montréal, through its end of doctoral studies scholarship and the excellence scholarship of the DIRO, as well as the support of CIRRELT through its doctoral excellence scholarship. Partial funding for this project has also been provided by the Natural Sciences and Engineering Research Council of Canada (NSERC), through its Discovery Grant program. We also gratefully acknowledge the support of Fonds de recherche du Québec through their infrastructure grants.

Results and views expressed in this publication are the sole responsibility of the authors and do not necessarily reflect those of CIRRELT.

Les résultats et opinions contenus dans cette publication ne reflètent pas nécessairement la position du CIRRELT et n'engagent pas sa responsabilité.

---

\* Corresponding author: Mohammad.Akhavan@cirrelt.ca

## 1. Introduction

*Network design* is a well-known and important class of problems in combinatorial optimization. *Multilayer network design* represents a special case of network design that has major applications in the fields of transportation (see, e.g., Cordeau et al., 2001; Zhu et al., 2014; Crainic et al., 2014) and telecommunications (see, e.g., Dahl et al., 1999; Knippel and Lardeux, 2007). Unlike a typical network design problem, in multilayer network design, there are several networks, each at a given layer. Each network has its nodes, potential arcs with (or without) limited capacities, and, possibly, commodities. The demands of commodities, if any, need to be routed from their origins to their destinations in each layer. To route the commodities, appropriate arcs have to be selected (or opened) by paying a fixed cost. A particular layer might not have any commodity, but still has to be designed to support the routing of other layers. At least one layer has commodities to route.

In multilayer network design, there are two types of coupling constraints between the layers, *flow connectivity* and *design connectivity* requirements. A common type of design connectivity requirement arises when each link in a layer can be selected only if some arcs (typically forming a path or a cycle) are opened in another layer. The flows in a layer might also be related to the flows of another layer, corresponding to flow connectivity requirements. For example, the amount of flow on each arc in a particular layer might be computed based on the flow on several arcs in another layer.

Connectivity requirements between layers might be either one-to-one or one-to-many. When each layer is supporting or is supported by only one other layer, the connectivity requirement is one-to-one while a one-to-many connectivity requirement exists when at least one of the layers is supporting or is supported by more than one layer. Note that, for two-layer network design problems, only the one-to-one connectivity is possible. In general, the objective is to find a minimum cost design and routing for all layers, while satisfying typical network design constraints in each layer, as well as coupling constraints between layers.

Multilayer network design is often, but not only, used to integrate decisions at the same planning level or different planning levels: strategic, tactical and operational. Solving the multilayer network design problem typically generates an optimal solution that cannot be obtained by solving sequentially each of the single-layer network design problems, thus yielding significant cost savings.

An example of such an integration can be found in railway freight planning, where cars have to be classified in groups called *blocks*. Then, blocks are grouped into services to make up trains moving blocks between terminals. Grouping cars into blocks avoids performing operations on each car individually in terminals, which reduces the number of operations to be performed. Zhu et al. (2014) addressed both problems of determining blocks (which block to be built) and selecting services in a single *integrated freight rail service network design problem*. They represent the problem using a three-layer network including car, block and service layers. Each layer consists of a *time-space network*, where the terminals (physical nodes) are duplicated over the time horizon to represent the time dependency. A node in such a network represents a terminal at a specific time, and each arc represents a transfer from a terminal at a given time to either the same terminal at another time or another terminal at another time. The service layer, includes moving and stop links of services. The block layer includes service section arcs (each corresponds to a chain of moving and stop arcs in the service layer) and block transfer arcs to move blocks between service sections. The car layer consists of block links (each corresponds to a chain of block transfer arcs and service sections in the block layer) and car arcs on which cars are moved

in each terminal. To open a service section arc in the block layer, a chain of moving and stop link should be open in the service layer. To select a projected block link in the car layer, a chain of block and projected service section links need to be open in the block layer.

Another example can be found in telecommunications, where one layer might be an internet (virtual) network whose arcs are supported by the arcs in an optical fiber (physical) layer. A chain of supporting arcs has to be opened in the physical layer to open an arc in the internet network. In this example, there is an integration of a strategic decision (physical network design) with a tactical one (virtual network design).

The applications of network design models and their solution techniques have been surveyed in Magnanti and Wong (1984), Minoux (1989), and Crainic (2000). In recent years, a growing number of applications of multilayer network design have appeared that are not covered in these surveys. To the best of our knowledge, the only survey paper is Kivelä et al. (2014), which does not cover extensively multilayer network design problems (only a few references for telecommunications applications are cited).

The contribution of this paper is threefold. First, we propose a classification of multilayer network design problems, which emphasizes the multilayer features, such as the number of layers and the type of coupling constraints between layers. Second, we synthesize the applications in transportation and telecommunications, as well as the methods used to solve multilayer network design problems. Third, we propose a general modeling framework that encompasses most multilayer network design problems found in the literature.

The paper is organized as follows. In Section 2, we propose a detailed definition of multilayer network design, as well as a general modeling framework. In Section 3, we present our classification of multilayer network design problems. Based on the classification and the applications currently proposed in the literature, we identify four existing classes of problems: 1) two-layer network design problems with design connectivity requirements; 2) two-layer network design problems with flow connectivity requirements; 3) three-layer network design problems with one-to-many flow connectivity requirements; 4) three-layer network design problems with one-to-one flow-design connectivity requirements. For each of these four classes of problems, a detailed survey is presented in Sections 4, 5, 6, and 7, respectively. In particular, we show how the proposed general modeling framework models most of the problems presented in the literature. In Section 8, we provide a survey on the proposed methods for solving multilayer network design problems. In Section 9, we summarize this work and we discuss future research directions.

## 2. Multilayer Network Design, Definition and Formulation

In this section, we first propose a definition of multilayer network design. We then propose a general modeling framework.

### 2.1. Definition

In network design, given a potential network (for simplicity, we assume all arcs are potential) that might have capacitated arcs, several commodities such as goods, data or people, have to be routed between different origin and destination points. A network has to be constructed by opening appropriate arcs between pairs of nodes to route the commodities. In addition to *flow costs*, *design costs* are associated to each arc. Flow costs are incurred when routing commodities on each arc, while a design cost is incurred when opening (or selecting) an arc. The problem is

to select the arcs such that the demands for the commodities can be routed on the constructed network, and arc capacities are respected. The designed network and the final routing must minimize the total cost.

In multilayer network design, instead of one network, there are several networks. Each network corresponds to a layer that consists of nodes and potential arcs. Several commodities might need to be routed in each layer to satisfy demands between origin and destination nodes. A network has to be designed in each layer to satisfy the demands for the commodities. Note that some layers might not have any commodity, but their arcs have to be opened to support the routing of the commodities in other layers. When there is only one layer that has commodities to route, we have a *multilayer single flow-type network design problem*. When we have commodities on more than one layer, we obtain a *multilayer multiple flow-type network design problem*.

In addition to flow and design costs, as well as flow capacities that might be associated with arcs, in multilayer network design, there are typically two types of coupling constraints between layers: *design connectivity* and *flow connectivity*. The first one, design connectivity, means that an arc opened in a given layer requires some arcs to be opened in another layer. If arc  $a$  of layer  $l'$  requires a set of arcs (for example a set including arcs  $b$  and  $c$ ) to be opened in layer  $l$ , then  $l'$  is said to be *supported* by  $l$ , and  $l$  is said to be *supporting* of  $l'$ . In addition,  $a$  is *supported* by  $b$  and  $c$ , while  $b$  and  $c$  are *supporting*  $a$ .

An illustration is given in Figure 1, where arcs 1 and 2 in layer  $l'$  are supported, respectively, by paths (3,4) and (5,6,4) in layer  $l$ . Therefore,  $l'$  is supported by  $l$ , and  $l$  is supporting  $l'$ . In this particular example, to use arc 1 in layer  $l'$ , all its supporting arcs in layer  $l$ , including arcs 3 and 4, have to be opened. For instance, in the integrated freight rail service network design problem, the design connectivity constraint consists of opening a chain of supporting services in the service layer to select the corresponding block in the block layer. In telecommunications, the design connectivity constraint forces an arc opened in the virtual network (the network supported by the physical layer) to be supported by a chain of physical arcs in the physical layer (the supporting network of the virtual layer). Note that, design connectivity requirements are not limited to the above examples. Another type arises when an arc in a layer requires at least one of the supporting arcs to be opened in another layer (for more details, see the next subsection where we describe different types of design connectivity requirements).

Based on the design connectivity constraints, we can define the *design capacity* constraints, a new concept in multilayer network design. The design capacity constraint of arc  $b$  in supporting layer  $l$  limits the number of selected arcs in supported layer  $l'$  for which arc  $b$  is the supporting arc. To clarify the definition of design capacity constraints, consider Figure 1. If the design capacity of arc 4 in layer  $l$  is equal to 1, then at most one of arcs 1 or 2 in layer  $l'$  can be opened in a feasible solution. For example, in the integrated freight rail service network design problem, a design capacity is defined for each service  $s$  and limits the number of selected blocks for which service  $s$  serves as the supporting service. In telecommunications, a design capacity constraint might be defined for each physical arc  $p$  to limit the number of opened virtual arcs for which arc  $p$  serves as a supporting arc.

The second type of connectivity constraints, flow connectivity, relates the flows between different layers. The simplest such constraint arises when the flow on arc  $b$  is equal to the summation of the flows on all arcs for which  $b$  is a supporting arc. In Figure 1, for example, the flow on arc 4 would be equal to the summation of the flows on arcs 1 and 2. Note that, with this particular type of flow connectivity requirements, when only one layer has commodities to route, the flows on other layers can be deduced from the flows on that single layer. Such a problem would be considered as a multilayer single flow-type network design problem, even though there

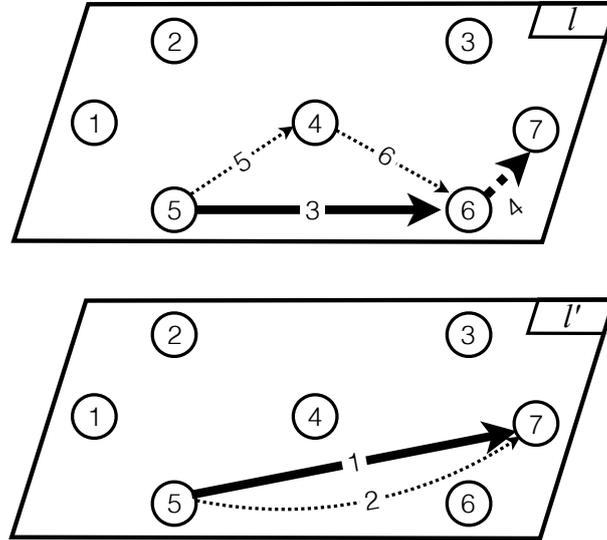


Figure 1: Multilayer network design example illustrating design connectivity constraints: arc 1 in layer  $l'$  is supported by the path made of arcs 3 and 4 in layer  $l$ , and arc 2 in layer  $l'$  is supported by the path made of arcs 5, 6, and 4

are flows on several layers. In the next subsection, we describe other types of flow connectivity constraints.

In some applications of multilayer network design, certain arcs of a layer can be independent of other layers. For example, in the integrated freight rail service network design problem, there are some arcs to move cars in each terminal that are not related to any arc of other layers.

Some problems introduced in the literature appear at first sight to be similar to the multilayer network design problem. These problems include the *multi-echelon network design problem* (Cordeau et al., 2006; Crainic et al., 2009), the *multilevel network design problem* (Balakrishnan et al., 1994; Costa et al., 2011), and the *hierarchical network design problem* (Obreque et al., 2010; Lin, 2010). There are two main differences between these problems and multilayer network design. First, in multilayer network design, each layer corresponds to a network including nodes, potential arcs, commodities to be routed on the designed network, while the concepts of echelon, level, and hierarchy do not necessarily correspond to a network with all these characteristics. The second main difference is the flow and design connectivity constraints between layers, which do not explicitly exist in the above problems.

## 2.2. Formulation

Given a set of layers  $L$  and a network  $G_l = (N_l, A_l)$  for each layer  $l \in L$ , where  $N_l$  and  $A_l$  are the sets of nodes and arcs of layer  $l \in L$ , respectively, we define  $u_{al}$  and  $v_{al}$  as the flow capacity and the design capacity of arc  $a \in A_l$  in layer  $l \in L$ . Let  $A_l^+(n)$  and  $A_l^-(n)$  represent the sets of outgoing and incoming arcs of node  $n \in N_l$ . A set of commodities  $K_l$  has to be routed through the network of layer  $l \in L$ . The set  $K_l$  might be empty, which means that there is no commodity to be routed in layer  $l$ . The amount of each commodity  $k \in K_l$  that must flow from its origin  $O(k) \in N_l$  to its destination  $D(k) \in N_l$  is  $d^k$ .

We denote by  $C$  the set of ordered pairs  $(l, l')$  such that  $l' \in L$  is a layer supported by  $l \in L$ . In other words,  $C$  contains the pairs of layers having a (design or flow) connectivity requirement between one another. Hence, we call  $C$  the set of connectivity requirement pairs. Let  $B_{l'}^{al}$  be the set of arcs in layer  $l'$  supported by arc  $a \in A_l$ . For example, in Figure 1, this set for arc 4 in layer  $l$  is  $\{1, 2\}$ . Let  $D_{bl'}^l$  be the set of arcs in layer  $l$  supporting arc  $b \in A_{l'}$ . In Figure 1, this set for arc 1 in layer  $l'$  is  $\{3, 4\}$ .

Two sets of decision variables are considered to formulate the problem, design and flow variables. The design variables could be binary or integer, depending on the particular application. When the decision is to open (select) or close (not to select) arc  $a \in A_l$  of layer  $l \in L$ , then the design variable  $y_{al}$  assumes binary values. When the goal is to determine the number of capacity units on each arc  $a \in A_l$  of layer  $l \in L$ , then the design variable  $y_{al}$  has integer values. The flow variables could take binary or continuous values depending on the problem. When the flow of each commodity has to be routed through a single path from its origin to its destination (non-bifurcated flows), then the flow variables take binary values. The variable  $x_{al}^k$  then indicates if commodity  $k \in K_l$  of layer  $l \in L$  uses arc  $a \in A_l$  or not. When the flow of each commodity can be distributed through several paths, then the flow variable  $x_{al}^k$  is continuous, representing the fraction of the demand of commodity  $k \in K_l$  of layer  $l \in L$  on arc  $a \in A_l$ . Sets  $X$  and  $Y$  define required side constraints, as well as the domains of the flow and design variables, respectively. Set  $(X, Y)_{ll'}$  defines the coupling constraints for each pair of layers  $(l, l') \in C$ , which captures some application-specific connectivity requirements. We present several possible coupling constraints later in this section.

We use notations  $\Psi(x)$  and  $\Phi(y)$  to represent the total routing cost function and the total design cost function, respectively. The proposed general multilayer network design formulation (MLND) can be stated as follows:

$$\min \quad \Psi(x) + \Phi(y) \quad (1)$$

subject to

$$\sum_{a \in A_l^+(n)} x_{al}^k - \sum_{a \in A_l^-(n)} x_{al}^k = w_n^k \quad \forall l \in L, \quad \forall n \in N_l, \quad \forall k \in K_l \quad (2)$$

$$\sum_{k \in K_l} d^k x_{al}^k \leq u_{al} y_{al} \quad \forall l \in L, \quad \forall a \in A_l \quad (3)$$

$$(x, y) \in (X, Y)_{ll'} \quad \forall (l, l') \in C \quad (4)$$

$$x \in X \quad (5)$$

$$y \in Y \quad (6)$$

The objective of the MLND model, (1), is to minimize the total routing and design cost. Constraints (2) are the usual *flow conservation equations*, ensuring that the demands are routed from their origins to their destinations in each layer, where  $w_n^k = 1$  if  $n = O(k)$ ,  $w_n^k = -1$  if  $n = D(k)$ , and 0 otherwise. The *flow capacity constraints* (3), ensure that the sum of the flows on each arc  $a \in A_l$  in layer  $l \in L$  does not exceed its flow capacity  $u_{al}$ .

Constraints (5) and (6) define side constraints and the domains of the decision variables. There are several side constraints that can be added to a network design problem, among which *design balance* and *budget constraints* are the most important ones. Design balance constraints arise when, at each node, the number of incoming opened arcs (representing, for example, re-

sources or vehicles) must be equal to the number of outgoing opened arcs. A budget constraint limits the cost for building the whole network to a total budget.

Constraints (4) are a set of *coupling constraints* that, for each connectivity requirement  $(l, l')$ , link together the domains of the decision variables  $(x, y)$  of layer  $l'$  to those of layer  $l$ . Several types of coupling constraints are encountered in the literature, depending on the applications.

In particular, *design capacity constraints* ensure that, for each arc  $a \in A_l$ , the number of selected arcs supported by  $a$  in layer  $l'$ , represented by set  $B_{l'}^{al}$ , does not exceed its design capacity  $v_{al}$ :

$$y_{al} \leq \sum_{b \in B_{l'}^{al}} y_{bl'} \leq v_{al} y_{al} \quad \forall (l, l') \in C, \quad \forall a \in A_l. \quad (7)$$

The left inequality of (7) ensures that if we open arc  $a$  in supporting layer  $l$ , then at least one of its supported arcs has to be opened in the supported layer  $l'$ .

A second type of coupling constraints is the *multilayer all-design linking constraints*:

$$y_{bl'} \leq y_{al} \quad \forall (l, l') \in C, \quad \forall a \in A_l, \quad \forall b \in B_{l'}^{al}. \quad (8)$$

These constraints simply mean that, to open arc  $b \in A_{l'}$ , all its supporting arcs have to be opened in all supporting layers  $l$ . Such constraints arise, for example, in the integrated freight rail service network design problem, where to open a block, all its supporting services have to be opened in the service layer.

A third type of coupling constraints is the *multilayer min-design linking constraints*:

$$y_{bl'} \leq \sum_{a \in D_{bl'}^l} y_{al} \quad \forall (l, l') \in C, \quad \forall b \in A_{l'}. \quad (9)$$

These constraints ensure that, for each arc  $b$  in a supported layer  $l'$ , at least one arc has to be opened in the supporting layer  $l \in L$ . Note that, constraints (8) imply (9), therefore, in general, one of them might be included in the formulation.

In addition to the above design connectivity coupling constraints, flow connectivity requirements between layers, whenever they are needed, might be added to the formulation. A first type of flow connectivity requirements is the *flow accumulation constraints*:

$$x_{al}^k = \sum_{b \in B_{bl'}^{al}} x_{bl'}^k \quad \forall (l, l') \in C, \quad \forall a \in A_l, \quad \forall k \in K. \quad (10)$$

These constraints simply mean that the flow on each arc  $a$  in layer  $l$  is equal to the flow on all the arcs in layer  $l'$  supported by arc  $a$ .

Note that, in some particular cases, constraints (10) might contradict flow conservation constraints (2). An example is when an arc in layer  $l$  supports two or more *reachable arcs* in layer  $l'$ . Two arcs  $(x_1, y_1)$  and  $(x_2, y_2)$  are said to be reachable if there is a path from  $x_1$  to  $y_2$  or from  $x_2$  to  $y_1$ . Consider Figure 2 as an example where arcs  $a$  and  $b$  are supported, respectively, by paths (1, 2) and (4, 6, 2, 7) in layer  $l$ . Suppose that there is a path between arcs  $a$  and  $b$  (the dashed arc in layer  $l'$ ), i.e., arcs  $a$  and  $b$  are reachable. Arc 2 in layer  $l$  supports both arcs  $a$  and  $b$  in layer  $l'$ . Suppose that a commodity with demand  $d$  has to be routed from node  $A$  to node  $B$  using arcs  $a$  and  $b$ , as well as the path between these two arcs (dashed arc). Based on equation (10), the flow on arc 2 is equal to the summation of the flows on arcs  $a$  and  $b$ , which is  $2d$ . If we consider the destination node of arc 2, its total incoming flow is  $2d$ , but its total outgoing flow (on arc 7) is

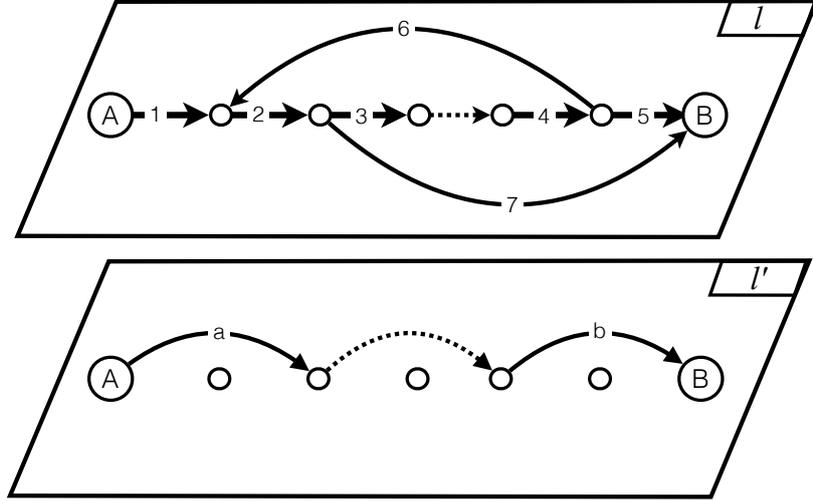


Figure 2: Multilayer network design example showing flow connectivity constraints contradict flow conservation constraints

*d.* So equations (10) contradict flow conservation constraints (10). Therefore, constraints (2) are not correct on general networks.

This issue does not arise in time-space networks, such as those used in Zhu et al. (2014). In this type of network, since the arcs are pointing to the next planning horizon, it is not possible that an arc supports two reachable arcs in another layer.

Another form of flow connectivity requirements might exist when flows are non-bifurcated, which means the flow of each commodity has to be routed through a single path from its origin to its destination. In this case, the following *non-bifurcated flow connectivity constraints*, might be added to the model:

$$\sum_{k \in K_{l'}} x_{b'l'}^k \leq \sum_{k \in K_l} x_{al}^k \quad \forall (l, l') \in C, \quad \forall a \in A_l, \quad \forall b \in B_{l'}^{al}. \quad (11)$$

For each arc  $a \in A_l$  and  $b \in B_{l'}^{al}$ , these constraints state that the flow of commodities  $K_{l'}$  can move on arc  $b$  in layer  $l'$  only if there is a flow of commodities  $K_l$  on arc  $a$  in layer  $l$ .

### 3. Multilayer Network Design Taxonomy

In this section, we propose a classification of multilayer network design problems, and we provide an overview of the existing literature, in light of the proposed classification.

#### 3.1. Multilayer Network Design Classification

Multilayer network design problems can be categorized into different classes based on three main dimensions. The first dimension is the number of layers. The second dimension is the

degree of connectivity between layers that includes *one-to-one connectivity* and *one-to-many connectivity*. A one-to-one connectivity exists when each layer is supporting or is supported by only one other layer. A one-to-many connectivity exists when at least one of the layers is supporting or is supported by more than one layer. Note that, for two-layer network design problems, the only possible degree of connectivity is the one-to-one connectivity.

The last dimension is the type of connectivity that includes design connectivity, flow connectivity, and *flow-design connectivity*. The last term is used when both types of connectivity requirements present at the same time. In the integrated freight rail service network design problem, for example, not only the designs of the layers are connected, but also the flow of each service is equal to the summation of the flows on its supported blocks. Figure 3 illustrates these three dimensions.

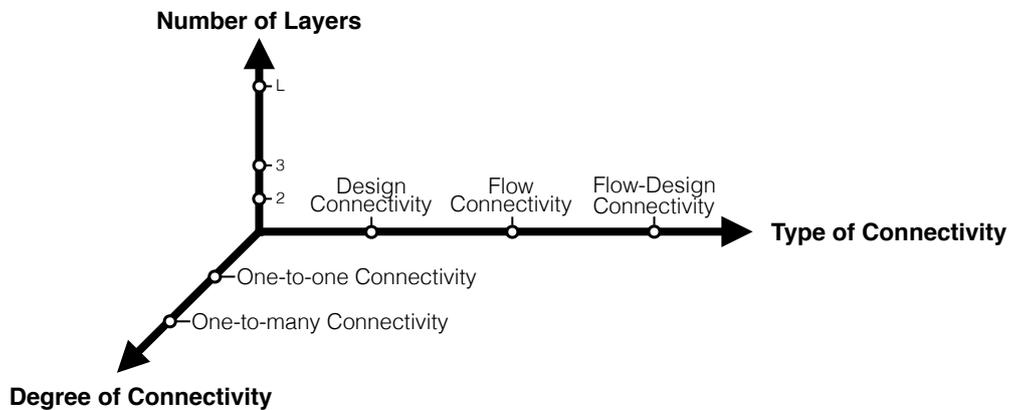


Figure 3: Classification dimensions of multilayer network design problems

### 3.2. Overview of the Literature

Given the proposed taxonomy, the existing literature fits into 4 classes: 1) two-layer network design problems with design connectivity; 2) two-layer network design problems with flow connectivity; 3) three-layer network design problems with one-to-many flow connectivity; 4) three-layer network design problems with one-to-one flow-design connectivity. There are some works in the literature on telecommunications applications that introduce multilayer network design models with  $L$  arbitrary layers (Orlowski and Wessály, 2004; Knippel and Lardeux, 2007), but they focus exclusively on two-layer applications.

The first category, two-layer network design problems with design connectivity, includes the *service network design with resource management* (Crainic et al., 2014, 2018) and most of the telecommunications applications (Dahl et al., 1999; Capone et al., 2007; Knippel and Lardeux, 2007; Fortz and Poss, 2009; Belotti et al., 2008; Koster et al., 2008; Orlowski, 2009; Raack and Koster, 2009; Mattia, 2012, 2013). The second one, two-layer network design problems with flow connectivity, consists of the *integrated crew scheduling and aircraft routing problem* (Cordeau et al., 2001; Cohn and Barnhart, 2003; Mercier et al., 2005; Mercier and Soumis, 2007; Shao et al., 2015; Salazar-González, 2014; Cacchiani and Salazar-González, 2016). The

third class, three-layer network design problems with one-to-many flow connectivity, includes the *integrated crew pairing and assignment problem* proposed in Zeighami and Soumis (2017). The only research work that falls into the fourth category, three-layer network design problems with one-to-one flow-design connectivity, is the *integrated rail freight service network design problem* proposed by Zhu et al. (2014).

Figure 4 summarizes all the proposed multilayer network design problems in the literature. This figure shows that the only paper that considers one-to-many connectivity is Zeighami and Soumis (2017). It also shows that almost all contributions in the literature consider two-layer network design problems with flow or design connectivity. The only paper that considers both flow and design connectivity requirements is Zhu et al. (2014).

		Type of Connectivity		
		Design Connectivity	Flow Connectivity	Flow-Design Connectivity
Number of Layers	2	<ul style="list-style-type: none"> <li>Dahl et al. (1999)</li> <li>Capone et al. (2007)</li> <li>Knippel &amp; Lardeux (2007)</li> <li>Belotti et al. (2008)</li> <li>Koster et al. (2008)</li> <li>Fortz &amp; Poss (2009)</li> <li>Orlowski (2009)</li> <li>Mattia (2012)</li> <li>Crainic et al. (2014)</li> <li>Crainic et al. (2018)</li> </ul>	<ul style="list-style-type: none"> <li>Cordeau et al. (2001)</li> <li>Cohn &amp; Barnhart (2003)</li> <li>Mercier et al. (2005)</li> <li>Mercier &amp; Soumis (2007)</li> <li>Shao et al. (2015)</li> </ul>	
	3		<div style="border: 1px solid black; padding: 5px; margin-bottom: 5px;">one-to-one</div> <div style="border: 1px solid black; padding: 5px;">one-to-many • Zeighami &amp; Soumis (2018)</div>	<div style="border: 1px solid black; padding: 5px; margin-bottom: 5px;">one-to-one • Zhu et al. (2014)</div> <div style="border: 1px solid black; padding: 5px;">one-to-many</div>
	L	<div style="border: 1px solid black; padding: 5px; margin-bottom: 5px;">one-to-one • Orlowski &amp; Wassaly (2004) • Knippel &amp; Lardeux (2007)</div> <div style="border: 1px solid black; padding: 5px;">one-to-many</div>		

Figure 4: Classification of existing multilayer network design problems in the literature

The next three sections present a comprehensive survey on the proposed multilayer network design models in the fields of transportation and telecommunications for each of the three main classes of problems identified above.

#### 4. Two-Layer Network Design Problems with Design Connectivity

In the following subsections, we review the service network design problem with resource management and the telecommunications applications, the two main classes of problems that fall into the category of two-layer network design problems with design connectivity.

#### 4.1. Service Network Design with Resource Management

Service network design with resource management is the only application in the literature of transportation that falls into the class of two-layer network design problems with design connectivity. In the following subsections, we first define the problem and review the existing literature, then we show how the proposed general modeling framework can model this problem.

##### 4.1.1. Problem Definition and Literature Review

Service network design models are broadly used in the field of transportation to formulate tactical planning problems. Most of these models assume the necessary resources (such as crews, power units, specific vehicles) are available at each terminal when needed. To overcome this simplification, researchers proposed models recognizing resource management aspects in service network design (see, e.g., Andersen et al. (2009a,b)). Crainic et al. (2014) enlarged the considered range of resource management issues. The required resources to perform services are considered to be assigned to the terminals to which they must ultimately return, and a limited number of resources are assigned to each terminal. The problem is to select services in the time-space network and to route the commodities on the selected services, while the services have to be supported by appropriate resource cycles in the resource layer, and the total cost of the designed network, resource assignment and routing has to be minimized.

To model the problem, the authors proposed a two-layer network including a time-space service network and a time-space resource cycle network. In the time-space service network, an arc is a service moving between terminals and times. In the time-space resource cycle network, an arc is defined as a resource cycle, which corresponds to a path of services from a terminal in time  $t$  to the same terminal in time  $t + TMAX$ , where  $TMAX$  is the maximum schedule length. There is a design connectivity constraint in the problem where each arc (a resource cycle) in the resource layer corresponds to a cycle of the supported services in the service layer. We illustrate these notions in Figure 5, where cycles  $r1$  and  $r2$  in the resource layer support the sets of services  $\{s1, s2, s3\}$  and  $\{s4, s5, s6\}$  in the service layer, respectively. Using the explained two-layer network, the authors proposed a cycle-based model to formulate the problem.

Crainic et al. (2014) assumed that there is only one type of resources and that the assignment of resources to terminals has been determined a priori. Crainic et al. (2018) extended this research in two ways: considering multiple types of resources, as well as the strategic decision of fleet acquisition and assignment. The problem has two layers and includes assignment or location decisions in the resource layer.

##### 4.1.2. MLND Formulation

To formulate the service network design problem with resource management with the general model presented in Section 2.2, we use the problem description in Crainic et al. (2014). We denote by  $K$  the set of commodities (single flow-type) and by  $d^k$  the demand for commodity  $k \in K$ . In this problem, we have a service layer ( $l = 1$ ) and a resource layer ( $l = 2$ ). The connectivity set  $C$  is defined as  $\{(2, 1)\}$ , meaning that the service layer is supported by the resource layer. To open an arc  $b \in A_1$  in the service layer, one of the supporting resource arcs in set  $D_{b1}^2$  should be opened in the resource layer. Let  $V$  be a set of terminals, then  $\theta_v$  is the set of resource arcs of layer 2 that depart from terminal  $v \in V$  during the scheduling length. We also denote by  $h_v$  the limit on the number of resources that depart at each terminal  $v \in V$ . Let  $f_{al}$  be the fixed cost of each arc  $a \in A_l, l \in L$  that has to be paid to open the corresponding service or resource arc. We also denote by  $c_{al}^k$  the routing cost of commodity  $k \in K$  on arc  $a \in A_l$ . For the service layer, the

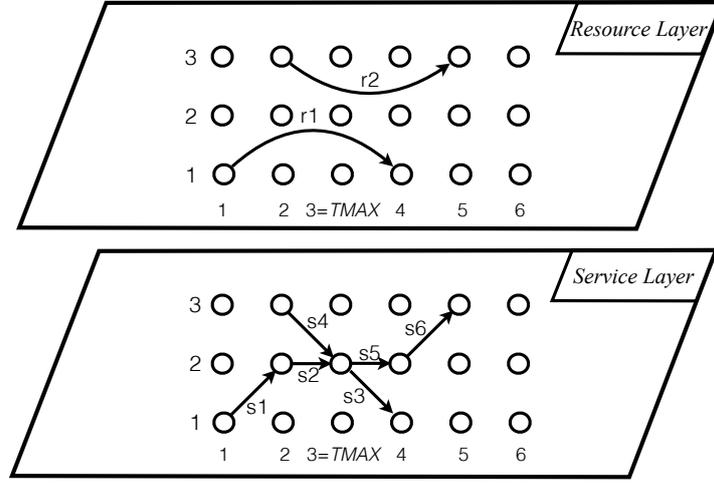


Figure 5: Resource cycles and their supported services in the service network design problem with resource management

decision variable  $x_{a1}^k$  determines the flow of each commodity  $k \in K$  on each arc  $a \in A_1$ . Let  $y_{a1}$  and  $y_{a2}$  be the design variables of service  $a \in A_1$  and resource  $a \in A_2$ , respectively. Then, the problem can be formulated as follows:

$$\min \sum_{k \in K} \sum_{a \in A_1} c_{a1}^k x_{a1}^k + \sum_{l \in L} \sum_{a \in A_1} f_{al} y_{al} \quad (12)$$

subject to

$$\sum_{a \in A_1^+(n)} x_{a1}^k - \sum_{a \in A_1^-(n)} x_{a1}^k = w_n^k \quad \forall n \in N_1, \quad \forall k \in K \quad (13)$$

$$\sum_{k \in K} d^k x_{a1}^k \leq u_{a1} y_{a1} \quad \forall l \in L, \quad \forall a \in A_1 \quad (14)$$

$$y_{b1} \leq \sum_{a \in D_{b1}^+} y_{a2} \quad \forall b \in A_1 \quad (15)$$

$$\sum_{a \in \theta_v} y_{a2} \leq h_v \quad \forall v \in V \quad (16)$$

$$x_{a1}^k \geq 0 \quad \forall a \in A_1, \quad \forall k \in K \quad (17)$$

$$y_{al} \in \{0, 1\} \quad \forall l \in L, \quad a \in A_1 \quad (18)$$

The objective function, (12), is to minimize the summation of the routing costs of the service layer and the design costs of the service and resource layers. Constraints (13) are the flow conservation equations ensuring the demands are satisfied in the service layer. *Service flow capacity constraints* (14) ensure that the total flow on each service arc is less than or equal to the flow capacity of the service and that the service must be open in order to route the commodities. *Service-resource coupling constraints* (15) show that, to open a service arc, at least one of the resource arcs should be open in the resource layer. Constraints (13), (14) and (15) are equivalent

to constraints (2), (3) and (9) of the general modeling framework, respectively. *Terminal resource capacity constraints* (16) are the side constraints that impose a limit on the number of resources of layer 2 that depart from terminal  $v \in V$  during the scheduling length. Constraints (17) and (18) define the domains of the decision variables.

#### 4.2. Telecommunications Applications

In this subsection, we first provide an overview of the multilayer network design problems in telecommunications and a survey on the related literature. Then, we describe how a typical telecommunications application can be modeled using the proposed general modeling framework.

##### 4.2.1. Problem Definition and Literature Review

In telecommunications applications, there are generally two layers: a *virtual layer* (also called *logical layer*) and a *physical layer* (or optical transport network). There is a set of identical nodes that are duplicated in both layers. These nodes might represent switch points. The nodes and the links can have several features: design cost, flow cost, *virtual flow capacity*, *physical design capacity*, and *node capacity*. The virtual flow capacity limits the flow of commodities on each logical link. The physical design capacity limits the number of logical links from the logical layer that can be supported by a physical link. The node capacity limits the number of virtual or physical links that can originate from, or end to, a particular node.

A set of commodities with specific demand quantities need to be routed in the logical layer. A network has to be designed in the logical layer to transfer the commodities and satisfy their demands. Several links have to be opened, or facilities have to be installed, between different pairs of nodes to design the network. Opening a link in the logical layer depends on opening a path in the physical layer. A typical example of two-layer network in telecommunications is an internet backbone network that has to be designed based on a physical fiber network. A chain of links (a path) has to be opened in the physical layer to establish a connection in the internet network. Figure 6 shows a simple example of a two-layer network in telecommunications applications. Link  $a_1$  in the virtual layer corresponds to links  $b_1$  and  $b_2$  in the physical layer, and link  $a_2$  corresponds to links  $b_3$ ,  $b_4$ , and  $b_5$ . To open or install facilities on a link in the logical layer, all corresponding links in the physical layer have to be opened or need to have appropriate facilities. For example, to open link  $a_2$ , all arcs  $b_3$ ,  $b_4$  and  $b_5$  have to be opened.

In telecommunications applications, the researchers mostly proposed two-layer networks. All telecommunications applications also fall into the design connectivity category where the design of each link in the virtual layer corresponds to the design of all corresponding links in the physical layer. To the best of our knowledge, the concept of layered networks in telecommunications dates back to Balakrishnan et al. (1991). Dahl et al. (1999) proposed a two-layer network for a telecommunications application. The problem is known as the *PIPE*, where the objective is to find a minimum cost pipe (virtual links) selection and routing while considering the design capacity of the physical links. Note that, in this problem, each demand has to be routed on a single virtual path (demands are not splittable). Therefore, the routing and the design variables are binary in the proposed formulation.

Capone et al. (2007) proposed a model for a two-layer network design problem with node capacity and *multicast traffic* demand where instead of point-to-point commodities, each commodity has an origin and multiple destinations. Therefore, a flow solution for each commodity is a tree, not a path.

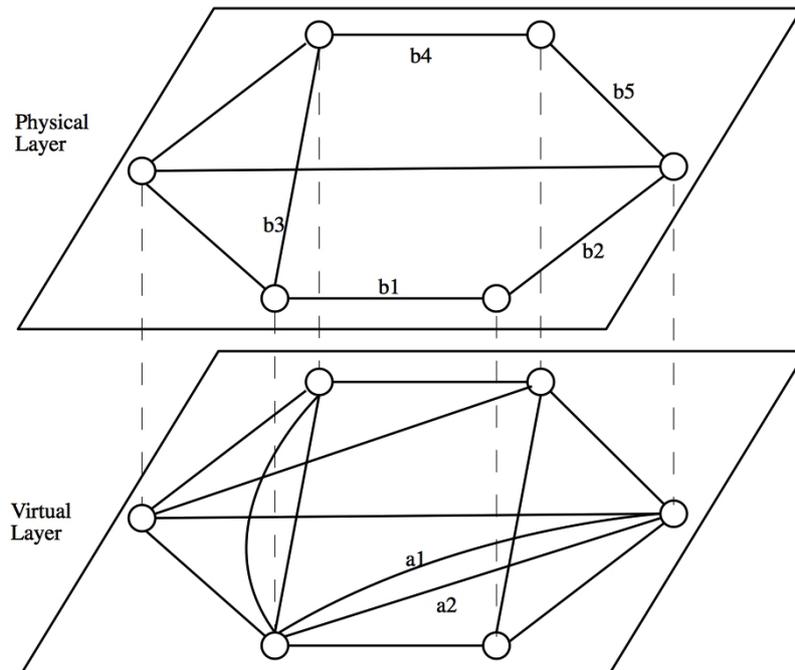


Figure 6: A two-layer telecommunications network

Knippel and Lardeux (2007) proposed a two-layer network design formulation, as well as a model with  $L$  arbitrary layers. The problem has fixed costs for the virtual and physical arcs, while no flow costs are considered. The model minimizes the total design cost of both layers. Parallel arcs are not used in the virtual layer; instead, the authors assumed each virtual flow capacity could be routed on several physical paths. Therefore, two types of continuous variables are introduced to determine the amount of each commodity on each logical path, and the amount of each installed logical traffic routed on each physical path. *Metric inequalities* are developed from the dual of the path-based formulation to represent the feasible space of capacity vectors.

Koster et al. (2008) proposed a formulation for a problem with a predefined set of logical links. The problem includes the selection of nodes and the *survivability requirements* against physical node and link failures. According to survivability requirements, a particular set of demands should be satisfied even if there is any single physical node or link failure. In the proposed formulation, the survivability requirements are presented using *survivability constraints*, where the demands are doubled, and the flow through an intermediate node is restricted to half of the demand value.

Mattia (2012) proposed a model that is similar to the one proposed in Knippel and Lardeux (2007) where the goal is to install minimum cost integer capacities on the links of both layers to route the commodities on them. In addition, survivability conditions are added to ensure that in every failure scenario the routing of the associated commodities must be guaranteed. Instead of adding the survivability constraints, several failure scenarios are defined in each of which just a

restricted number of links are available. For each failure scenario, a two-layer network is defined containing only the available links, with variables defined for each scenario.

There are other contributions in the literature on telecommunications applications that focus mostly on solution methods (Fortz and Poss, 2009; Orłowski, 2009; Orłowski et al., 2010; Raack and Koster, 2009; Mattia, 2013; Belotti et al., 2008). We review these papers in Section 8.

#### 4.2.2. MLND Formulation

We use the problem description in Dahl et al. (1999) as a representative of a telecommunications application that can be formulated using the general modeling framework presented in Section 2.2. We denote by  $L = \{1, 2\}$  the set of layers including the virtual layer ( $l = 1$ ) and the physical layer ( $l = 2$ ). Let  $C = \{(2, 1)\}$  be the set of connectivity requirements, where the ordered pair  $(2, 1)$  means that the physical layer is supporting the design of the virtual layer. We denote by  $K$  the set of single flow-type commodities to be routed on the virtual layer and by  $d^k$  the demand of each commodity  $k \in K$ . Let  $c_{a1}^k$  be the flow cost of routing one unit of commodity  $k \in K$  on arc  $a \in A_1$ , and  $f_{al}$  be the fixed cost of opening an arc  $a \in A_l, l \in L$ . The binary flow variable  $x_{a1}^k$  determines if the demand of commodity  $k \in K$  flows on the virtual arc  $a \in A_1$ . The binary design variables  $y_{a1}$  and  $y_{a2}$  determine, respectively, if a virtual arc  $a \in A_1$  and if a physical arc  $a \in A_2$  is open. Using the above notation, the problem can be formulated as follows:

$$\min \sum_{k \in K} \sum_{a \in A_1} c_{a1}^k x_{a1}^k + \sum_{l \in L} \sum_{a \in A_l} f_{al} y_{al} \quad (19)$$

subject to

$$\sum_{a \in A_1^+(n)} x_{a1}^k - \sum_{a \in A_1^-(n)} x_{a1}^k = w_n^k \quad \forall n \in N_1, \quad \forall k \in K \quad (20)$$

$$\sum_{k \in K} d^k x_{a1}^k \leq u_{a1} y_{a1} \quad \forall l \in L, \quad \forall a \in A_1 \quad (21)$$

$$y_{a2} \leq \sum_{b \in B_1^{a2}} y_{b1} \leq v_{a2} y_{a2} \quad \forall a \in A_2 \quad (22)$$

$$x_{a1}^k \in \{0, 1\} \quad \forall a \in A_1, \quad \forall k \in K \quad (23)$$

$$y_{al} \in \{0, 1\} \quad \forall a \in A_l \quad (24)$$

The objective function (19) is to minimize the total cost including the total routing cost of the virtual layer and the summation of the design costs of the virtual and physical layers. Constraints (20) are the flow conservation equations that ensure the demands are satisfied in the virtual layer. Flow capacity constraints (21) are imposed for the virtual layer. Design capacity constraints (22) are coupling constraints ensuring that, to open an arc in the virtual layer the supporting arcs in the physical layer should be open and that the maximum number of selected virtual arcs is limited to the design capacity of the corresponding physical arc. These constraints correspond to the design capacity constraints of the general modeling framework (7). Constraints (23) and (24) define the feasible domains of the decision variables. In telecommunications applications, the flow variables are binary, to ensure that the flow of each commodity follows a single path from the origin to the destination. Note that, when the links are undirected in a telecommunications application, the problem can still be modeled using the proposed formulation by replacing an undirected link with two directed arcs.

## 5. Two-Layer Network Design Problems with Flow Connectivity

The integrated crew scheduling and aircraft routing problem is, to the best of our knowledge, the only application that falls into the category of two-layer network design problems with flow connectivity. In Subsection 5.1, we provide an overview of the problem, and we review the literature on this topic. In Subsection 5.2, we explain how the model we propose in Section 2.2 can formulate this application.

### 5.1. Integrated Crew Scheduling and Aircraft Routing: Literature Review

The first step in the airline planning process is *flight scheduling* to define the origin and the destination, as well as the departure and arrival times, for each flight leg to be flown during a given period. The next step is to assign an aircraft type to each flight leg to maximize the profit, which is called the *fleet assignment problem*. Then, for each aircraft type, the *aircraft routing problem* is solved to determine the sequence of flight legs that have to be covered by each aircraft. In this problem, each flight leg has to be covered exactly once while ensuring aircraft maintenance requirements. The next step is called the *crew scheduling problem*, which consists in two steps: *crew pairings* followed by *crew assignment*. A crew pairing is a sequence of duty and rest periods starting and ending at the same location called crew base. A duty period is a sequence of flight legs separated by short rest periods. The duties are also separated by long (overnight) rest periods. The crew assignment is to build monthly scheduling out of generated pairings for each crew member.

Traditionally, most airlines use a sequential procedure to solve these problems. The sequential procedure reduces the complexity of the problem, but might result in a solution far from the global optimum of the integrated problem. The integrated crew scheduling and aircraft routing problem is an attempt to handle this issue. The problem is defined on a two-layer network including 1) an aircraft routing network, and 2) a crew scheduling network. In both networks, each node corresponds to a flight leg, and the arcs represent the connections between two legs. The integrated problem is to find the minimum total cost of aircraft and crew routing (one path for each aircraft and one path for each crew), while the two following conditions are satisfied: 1) each flight leg has to be covered only once by a crew and only once by an aircraft, and 2) if a connection time for a link is too short (*short time links*) then the corresponding legs can be covered by the same crew only if both legs are covered by the same aircraft; otherwise, the connection time is insufficient for the crew. The second condition corresponds to the second type of flow connectivity requirements shown in constraints (11).

Cordeau et al. (2001) proposed a path-based formulation for the integrated crew scheduling and aircraft routing problem. Cohn and Barnhart (2003) contributed to the literature on the integrated aircraft routing and crew scheduling problem by proposing an extended crew pairing formulation. In the proposed formulation, the aircraft routing variables represent a complete solution of a routing problem. Mercier et al. (2005) improved the integrated approach of Cordeau et al. (2001) by introducing *restricted connection arcs* in addition to short time connections. A connection is restricted if the connection time is larger than a minimum threshold, but it is still smaller than another given threshold. If the two legs of such a connection are covered by the same crew, a penalty is imposed in the objective function if both legs of this connection are not covered by the same aircraft. Mercier and Soumis (2007) extended the integrated approach by adding the *flight re-timing* feature, where flight legs have different possible departure times among which the best one has to be selected to minimize the cost.

Shao et al. (2015) integrated the fleet assignment problem to the integrated crew scheduling and aircraft routing problem. This new integration adds the decision of assigning a proper fleet type to each flight leg (node) of the network.

Salazar-González (2014) proposed an arc-based formulation where both aircraft routes and crew pairs are presented using arc-based variables. The main disadvantage of the proposed formulation is that a large number of inequalities need to be added to avoid infeasible crew routes. Two alternative formulations are also proposed in Cacchiani and Salazar-González (2016) including a path-based model as in Cordeau et al. (2001) and an arc-path-based model using arc-based variables and path-based variables to represent aircraft routes and crew pairings, respectively.

### 5.2. MLND Formulation for Integrated Crew Scheduling and Aircraft Routing

Using the description of Cordeau et al. (2001) as a representative of integrated crew scheduling and aircraft routing problems, we describe how to formulate the problem using the general modeling framework presented in Section 2.2. Let  $L = \{1, 2\}$  be the set of layers including a crew layer ( $l = 1$ ) and an aircraft layer ( $l = 2$ ). The nodes in both layers are the flight legs, while the arcs are the crew and aircraft connections, respectively, in the crew and aircraft layers. A set of crews ( $K_1$ ) and aircrafts ( $K_2$ ) should be routed on the crew and the aircraft layers, respectively. Note that the flows are non-bifurcated. We partition the nodes of each layer  $l \in L$  into  $N_l^O = \{n \in N_l \mid \exists k \in K_l, n = O(k)\}$ ,  $N_l^D = \{n \in N_l \mid \exists k \in K_l, n = D(k)\}$  and  $N_l^I = N_l \setminus N_l^O \cup N_l^D$ . Binary decision variable  $x_{a_1}^k$  and  $x_{a_2}^k$  determine, respectively, if crew  $k \in K_1$  uses arc  $a \in A_1$  and if aircraft  $k \in K_2$  uses arc  $a \in A_2$ . The set  $C$  is defined as  $\{(2, 1)\}$  representing the coupling constraints which indicates the aircraft layer supports the crew layer. Using the described notation, the problem can be formulated as:

$$\min \Psi(x) \quad (25)$$

$$\sum_{a \in A_1^+(n)} x_{a_1}^k - \sum_{a \in A_1^-(n)} x_{a_1}^k = w_n^k \quad \forall n \in N_1, \quad \forall k \in K_1 \quad (26)$$

$$\sum_{a \in A_2^+(n)} x_{a_2}^k - \sum_{a \in A_2^-(n)} x_{a_2}^k = w_n^k \quad \forall n \in N_2, \quad \forall k \in K_2 \quad (27)$$

$$\sum_{k \in K_1} x_{b_1}^k \leq \sum_{k \in K_2} x_{a_2}^k \quad \forall a \in A_2, \quad \forall b \in B_1^{a_2} \quad (28)$$

$$\sum_{a \in A_1^+(n)} \sum_{k \in K_1} x_{a_1}^k = 1 \quad \forall l \in L, \quad \forall n \in N_l^O \cup N_l^I \quad (29)$$

$$\sum_{a \in A_l^+(n)} \sum_{k \in K_l} x_{a_l}^k = 1 \quad \forall l \in L, \quad \forall n \in N_l^D \cup N_l^I \quad (30)$$

$$x_{a_l}^k \in \{0, 1\} \quad \forall a \in A_l, \quad \forall k \in K \quad (31)$$

The objective function (25) minimizes the total routing costs on both layers. In airline applications, since the objective is typically a non-linear function of arc-based flow variables, researchers usually propose path-based formulations for which the objective function is linear. Crew flow conservation equations (26) and aircraft flow conservation equations (26) guarantee, respectively, the routing of the flows of the crews and aircrafts on the crew and aircraft layers. Constraints (28) ensure that a crew does not change aircraft when the connection time is

too short. These constraints correspond to the flow connectivity inequalities (10) of the general modeling framework. Constraints (29) and (30) are the side constraints ensuring that a flight leg is covered by exactly one crew and one aircraft. Constraints (31) define the domain of the decision variables.

## 6. Three-Layer Network Design Problems with One-to-Many Flow Connectivity

To the best of our knowledge, the only paper that considers one-to-many connectivity requirements is Zeighami and Soumis (2017). In the following subsections, we describe the integrated crew pairing and assignment problem, and then we explain how the MLND framework can formulate this problem.

### 6.1. Integrated Crew Pairing and Assignment: Literature Review

The crew scheduling problem constructs individual schedules for a set of crew members. Because of its complexity, this problem is usually solved in two steps: crew pairing followed by crew assignment. The sequential procedure reduces the complexity of the problem but might result in a solution far from the global optimum of the integrated problem since the schedule constraints and objectives are not taken into account during the construction of the pairings.

Zeighami and Soumis (2017) proposed an integrated crew pairing and personalized assignment problem for a given set of pilots and copilots. They considered a set of vacation requests (VRs) for each pilot and copilot each month. The problem is defined on a three-layer network including 1) a crew pairing network, 2) a crew pilot assignment network, and 3) a crew copilot assignment network. In the crew pairing network, each node corresponds to a departure and arrival station of the flights. The arcs represent the flights and the connections between the flights. In the pilot (copilot) assignment network, each node corresponds to the start and end of pairings, and the arcs represent the pairings, the connections between pairings, and the vacations of the pilots (copilots). The objective function aims for a trade-off between maximizing the number of satisfied VRs and minimizing the total cost of the pairings (one path for each pairing and one path for each crew pilot (copilot) assignment), while two main conditions are satisfied: 1) each flight is covered by exactly one pairing, and 2) each pairing is covered by exactly one pilot and one copilot.

### 6.2. MLND Formulation for Integrated Crew Pairing and Assignment

To model the problem using the general modeling framework presented in Section 2.2, we define the following notation. Let  $L = \{1, 2, 3\}$  be the set of layers including the crew pairing ( $l = 1$ ), pilot assignment ( $l = 2$ ), and copilot assignment ( $l = 3$ ) layers.  $G_l = (N_l, A_l)$  defines the network of each layer  $l \in L$ . In the crew pairing layer, we partition the arcs into the sets of flight arcs  $A_1^f$  and connection arcs  $A_1^c$ . We assume that all the potential pairing arcs exist in the pilot and copilot assignment layers. There are a set of pilot  $K_2$  and copilots  $K_3$  that need to be routed in the pilot and copilot assignment layers, respectively.  $O(k) \in N_l$  and  $D(k) \in N_l$  are, respectively, the origin and the destination of each crew  $k \in K_l$  in layers  $l \in \{2, 3\}$ .

We denote by  $C = \{(1, 2), (1, 3)\}$  the set of connectivity requirements, where (1, 2) and (1, 3) mean that, respectively, the pilot and copilot assignment layers are supported by the crew pairing layer. To use a pairing arc in the assignment layer, all the corresponding arcs need to be selected in the crew pairing layer. Let  $B_2^{a_1}$  and  $B_3^{a_1}$  be the sets of arcs (pairings), respectively, in the pilot and copilot layers supported by arc  $a \in A_1$  in the pairing layer. Binary flow variable  $y_{a_1}$

determines whether arc  $a \in A_1$  is selected or not. Binary flow variables  $x_{a_2}^k$  and  $x_{a_3}^k$  determine, respectively, whether or not pilot  $k \in K_2$  and copilot  $k \in K_3$  selects arc  $a \in A_2$  and  $a \in A_3$ . Using this notation, the problem is formulated as follows:

$$\min \Psi(x, y) \quad (32)$$

$$\sum_{a \in A_2^+(n)} x_{a_2}^k - \sum_{a \in A_2^-(n)} x_{a_2}^k = w_n^k \quad \forall n \in N_2, \quad \forall k \in K_2 \quad (33)$$

$$\sum_{a \in A_3^+(n)} x_{a_3}^k - \sum_{a \in A_3^-(n)} x_{a_3}^k = w_n^k \quad \forall n \in N_3, \quad \forall k \in K_3 \quad (34)$$

$$\sum_{k \in K_2} x_{b_2}^k \leq y_{a_1} \quad \forall a \in A_1, \quad \forall b \in B_2^{a_1} \quad (35)$$

$$\sum_{k \in K_3} x_{b_3}^k \leq y_{a_1} \quad \forall a \in A_1, \quad \forall b \in B_3^{a_1} \quad (36)$$

$$\sum_{b \in B_2^{a_1}} \sum_{k \in K_2} x_{b_2}^k = 1 \quad \forall a \in A_1^f \quad (37)$$

$$\sum_{b \in B_3^{a_1}} \sum_{k \in K_3} x_{b_3}^k = 1 \quad \forall a \in A_1^f \quad (38)$$

$$x_{a_l}^k \in \{0, 1\} \quad \forall l \in \{2, 3\}, \quad \forall a \in A_l, \quad \forall k \in K_l \quad (39)$$

$$y_{a_1} \in \{0, 1\} \quad \forall a \in A_1 \quad (40)$$

$$x \in X \quad (41)$$

The objective function (32) minimizes the total routing and design costs on the three layers. Pilot flow conservation equations (33) guarantee the routing of each pilot  $k \in K_2$  in the second layer. The same type of flow conservation constraints (34) exist for the copilot layer. In this equations  $w_n^k = 1$  if  $n = O(k)$ ,  $w_n^k = -1$  if  $n = D(k)$ , and 0 otherwise. Constraints (35) and (36) are the coupling constraints ensuring the flow connectivity between layers. Constraints (37) and (38) are the covering constraints ensuring each flight arc is covered by exactly one pairing. Constraints (39) and (40) define the domain of the decision variables. Constraints (41) are side constraints including the restrictions corresponding to the VRs.

## 7. Three-Layer Network Design Problems with One-to-One Flow-Design Connectivity

The only paper that applies both flow and design connectivity requirements is Zhu et al. (2014). In the following subsections, we first describe the problem proposed in Zhu et al. (2014). Then, in the second subsection, we explain how the MLND framework can model this problem.

### 7.1. Integrated Rail Freight Service Network Design: Literature Review

Zhu et al. (2014) proposed a three-layer network to model a problem in rail freight transportation planning where typically a double consolidation policy is performed. First, cars are grouped into so-called blocks, and then the blocks are grouped into services to make up trains. Cars that are in a terminal at the same time can be sorted and arranged into a block. This process is called blocking, and its goal is to reduce operations in terminals by moving blocks instead of each car

individually. A block is then a unit that will be transferred between terminals using a sequence of services until it reaches its destination. At the block destination, the block is broken down, and the cars that arrived at their destinations are delivered, while the cars that did not reach their destinations are grouped into other blocks. The process of arranging the blocks into services is known as train makeup. The next step is to select the services and define their frequencies.

In Zhu et al. (2014), the blocking and service selection problems are considered together. To do so, a three-layer network including car network, block network, and service network is used. The network of each layer is a time-space network. The first layer, the car layer, consists of *car-waiting arcs*, *classification arcs* and *car-holding arcs*. Car-waiting arcs show the waiting of cars between two time periods in one terminal. Classification arcs show the classification process of cars between two time periods in one terminal. Car-holding arcs show waiting process of the classified cars between two time periods in one terminal. The car-waiting, classification, and car-holding arcs are not related to any path of the block layer. The second layer, the block layer, includes block arcs from the origins to the destinations of the blocks to support car movements. The third layer, the service layer, includes service arcs to support block movements.

In the car layer, the flows of commodities are moved via the car arcs and the projected block arcs from the block layer. A chain of services has to be opened in the service layer (design connectivity) to open a block arc. The flow of each service is equal to the summation of the flows on all its supported blocks (flow connectivity). The problem is to find a minimum cost blocking and service design, and flow routing of the cars, while considering flow capacity of the blocks and the services, blocking capacity of each terminal, and flow and design connectivity requirements between the layers.

### 7.2. MLND Formulation for Integrated Rail Freight Service Network Design

We define the following notation to model the problem using the general modeling framework presented in Section 2.2. Let  $K$  be a set of single flow-type commodities. There are three layers, the car layer ( $l = 1$ ), the block layer ( $l = 2$ ) and the service layer ( $l = 3$ ). The car layer includes car arcs and block arcs projected from the block layer. The block layer includes block holding, transfer and moving arcs. The service layer consists of service waiting and moving arcs. We denote by  $C = \{(3, 2), (2, 1)\}$  the set of connectivity requirements, where  $(2, 1)$  means that the block layer is supporting the car layer, and  $(3, 2)$  indicates that the service layer is supporting the block layer. In the car layer, each projected block arc is supported by a path of block holding, transfer and moving arcs in the block layer. In the block layer, a block moving arc is supported by a chain of service moving and waiting arcs of the service layer. For each layer  $l \in L$ , continuous flow variable  $x_{al}^k$  determines the flow of commodity  $k \in K$  on arc  $a \in A_l$ . For the car layer, the binary design variable  $y_{a1}$  determines the selection of a car arc or a projected block arc  $a \in A_1$ . In the block ( $l = 2$ ) and service ( $l = 3$ ) layers, the binary design variable  $y_{al}$  is 1 if arc  $a \in A_l$  is selected. We denote by  $T$ ,  $V$  and  $E$  the set of time periods, yards, and track segments, respectively. Let  $H(v, t)$  be the set of blocks built simultaneously at yard  $v \in V$ , and let  $S(e, t)$  be the set of services moved simultaneously at track segment  $e \in E$ . We denote by  $h_v$  and  $s_e$ , respectively, the maximum number of blocks and services that can be built simultaneously at each yard  $v \in V$  and track segment  $e \in E$ . Using the described notation, the problem is formulated as follows:

$$\min \sum_{l \in L} \sum_{k \in K} \sum_{a \in A_l} c_{al}^k x_{al}^k + \sum_{l \in L} \sum_{a \in A_l} f_{al} y_{al} \quad (42)$$

$$\sum_{a \in A_1^+(n)} x_{a1}^k - \sum_{a \in A_1^-(n)} x_{a1}^k = w_n^k \quad \forall n \in N_1, \quad \forall k \in K \quad (43)$$

$$x_{al}^k = \sum_{b \in B_p^{al}} x_{bl}^k \quad \forall (l, l') \in C, \quad \forall a \in A_l, \quad \forall k \in K. \quad (44)$$

$$\sum_{k \in K_l} d^k x_{al}^k \leq u_{al} y_{al} \quad \forall l \in L, \quad \forall a \in A_l \quad (45)$$

$$y_{a3} \leq \sum_{b \in B_2^{a3}} y_{b2} \leq v_{a3} y_{a3} \quad \forall a \in A_3 \quad (46)$$

$$y_{bl} \leq y_{al} \quad \forall (l, l') \in C, \quad \forall a \in A_l, \quad \forall b \in B_p^{al}. \quad (47)$$

$$\sum_{a \in H(v,t)} y_{a2} \leq h_v \quad \forall v \in V, \quad \forall t \in T \quad (48)$$

$$\sum_{a \in S(e,t)} y_{a3} \leq s_e \quad \forall e \in E, \quad \forall t \in T \quad (49)$$

$$x_{al}^k \geq 0 \quad \forall l \in L, \quad \forall a \in A_l, \quad \forall k \in K \quad (50)$$

$$y_{al} \in \{0, 1\} \quad \forall l \in L, \quad \forall a \in A_l \quad (51)$$

The objective function (42) minimizes the total routing and design costs on the three layers. Constraints (43) guarantee that the demands are routed on the car layer, while constraints (44) compute the flow on each arc of the block and service layers based on the corresponding arcs of the car layer. Flow capacity constraints (45) ensure that the flow on each arc is less than or equal to the capacity and that the arc should be opened in order to route the commodities. Constraints (47) ensure that, to open an arc in the car layer, all the corresponding block arcs must be opened in the block layer, and that, to open an arc in the block layer, all the corresponding arcs must be opened in the service layer. Design capacity constraints (46) limit the number of blocks that can be moved on the corresponding arc of the service layer. These constraints correspond to the design capacity constraints (7) of the general modeling framework. Constraints (48) and (49) are the side constraints that limit, respectively, the number of blocks and services to be created at each yard and track segment.

## 8. Solution Approaches for Multilayer Network Design Problems

In this section, we summarize the solution methods proposed in the literature for multilayer network design problems. The solution methods proposed in the literature can be classified as 1) exact solution methods (Dahl et al., 1999; Knippel and Lardeux, 2007; Fortz and Poss, 2009; Koster et al., 2008; Raack and Koster, 2009; Mattia, 2012, 2013; Cordeau et al., 2001; Mercier et al., 2005; Shao et al., 2015; Cacchiani and Salazar-González, 2016), and 2) heuristic solution methods (Orlowski, 2009; Orlowski et al., 2010; Crainic et al., 2014, 2018; Capone et al., 2007; Belotti et al., 2008; Salazar-González, 2014; Zhu et al., 2014). In the two following subsections, we review, respectively, the exact and heuristic methods proposed for solving multilayer network design problems.

### 8.1. Exact Solution Methods

In exact solution methods, researchers mostly focus on either embedding Benders decomposition and classical network design cuts into branch-and-bound or combining Benders decomposition and column generation. In the two following subsections, we review: 1) cutting plane and Benders decomposition methods (Dahl et al., 1999; Knippel and Lardeux, 2007; Fortz and Poss, 2009; Koster et al., 2008; Orłowski, 2009; Orłowski et al., 2010; Raack and Koster, 2009; Mattia, 2012, 2013), and 2) combined column generation and Benders decomposition methods (Cordeau et al., 2001; Mercier et al., 2005; Shao et al., 2015; Cacchiani and Salazar-González, 2016).

#### 8.1.1. Cutting Plane and Benders Decomposition Methods

Dahl et al. (1999) used a branch-and-cut algorithm for a two-layer telecommunications network design problem. At each node of the branch-and-bound tree, several valid inequalities are added to the linear programming (LP) relaxation. A variable fixing heuristic is called when the branch-and-cut algorithm does not find any violated inequality. Several instances derived from a real-world application are used to test the algorithm. The algorithm is able to solve to optimality all instances with no fixed costs, without making any branching (these instances are solved at the root). For the instances with positive fixed costs, the algorithm could not solve the problems to optimality but found solutions within an average optimality gap of 9%.

Knippel and Lardeux (2007) proposed a Benders decomposition algorithm for a two-layer telecommunications network design problem with no flow cost where the master problem handles the design variables, and two subproblems determine the value of the flow variables. The algorithm first solves the master problem as an integer program with no cuts. Then, the algorithm checks the feasibility of the obtained solution by solving two subproblems, one for the logical layer and the other for the physical layer. If the solution is infeasible, then the corresponding cuts are added to the master problem. The solution time of the master problem is larger than that of the subproblems because of the integrality conditions on the master problem variables. To reduce the solution time, different approaches for the generation of cuts are proposed.

Fortz and Poss (2009) improved the Benders decomposition method proposed in Knippel and Lardeux (2007). The idea is to embed the Benders decomposition into a branch-and-cut algorithm. In this way, instead of solving the master problem as an integer program, the algorithm solves the LP relaxation of the master problem at each node of the branch-and-bound tree. If the solution is integral, it adds the corresponding Benders cuts by solving the subproblem. If the solution is not integral, it generates branches in the tree and adds the branching constraints to the master problem. This way, the solution time is significantly reduced compared to that of the traditional Benders decomposition approach. To avoid generating too many infeasible nodes, cuts are added a priori to the master problem by solving its LP relaxation with the cutting plane algorithm proposed by Knippel and Lardeux (2007). The results are compared with the cutting plane approaches of Knippel and Lardeux (2007), and with CPLEX. The results show that the algorithm outperforms CPLEX, and that the proposed branch-and-cut algorithm is always faster than the Benders decomposition approaches.

Koster et al. (2008) proposed a cutting plane method embedded in a branch-and-bound framework for the telecommunications application with survivability requirements against physical node and link failures. In the proposed cutting plane method, the authors used the cuts that were applied before in the literature on the single layer network design problem. The algorithm is tested on three different sets of instances. In the case of unprotected demands (no survivability

requirement), the cutting plane algorithm significantly improves the LP relaxation lower bounds. In the case of protected demand, where the size of the problem dramatically increases compared to the unprotected case, the cutting plane algorithm only slightly improves the LP relaxation lower bounds.

Raack and Koster (2009) studied a packing problem derived from a two-layer network design problem. The authors proved the NP-hardness of the problem and defined two classes of facet-defining inequalities that generalize the well-known cutset inequalities to two-layer network design.

Mattia (2012) proposed a branch-and-cut scheme using metric inequalities for the same problem as in Knippel and Lardeux (2007), but with survivability requirements. Further work on the same problem can be found in Mattia (2013), where the polyhedron of a two-layer network design formulation is considered. The results are extended to multilayer network design models.

#### 8.1.2. Combined Column Generation and Benders Decomposition Methods

Orlowski (2009) proposed several techniques combining Benders decomposition and column generation embedded into a branch-and-cut framework for a multilayer network design problem with survivability requirements. The solution methods are based on the approaches proposed in Fortz and Poss (2009) and Koster et al. (2008). Similar to Fortz and Poss (2009), metric inequalities are used to create a Benders master problem by projecting out the flow variables to a subproblem. The LP relaxation of the Benders master problem is solved at each node of the branch-and-cut tree. Whenever an integer solution is found, the algorithm checks the feasibility in a routing subproblem, and adds cuts to the master problem, if the design is not feasible according to the routing constraints. The cuts proposed in Koster et al. (2008) are generated, along with cuts derived from survivability constraints. Several primal heuristics are proposed to improve the feasible integer solutions. A column generation approach is developed to generate flow variables dynamically in the large-scale routing subproblems. The algorithm could find feasible solutions and lower bounds for large-scale instances that could not be solved by a commercial solver. However, for large and dense instances, the obtained feasible solutions are still far from optimality (57% and 28% optimality gap on average for the instances with and without the survivability conditions, respectively).

Orlowski et al. (2010) combined the approaches in Orlowski (2009) and Koster et al. (2008). In addition to the cutting plane approach presented in Koster et al. (2008), the heuristic methods presented in Orlowski (2009) were also used. The heuristics are called at various places of the branch-and-cut tree. The algorithm is tested on several real-world industrial instances.

Cordeau et al. (2001) proposed a Benders decomposition approach for the integrated crew scheduling and aircraft routing problem. The classical solution method is to use branch-and-bound where, at each node, the LP relaxation is solved using a column generation method. However, in the computational experiments, the authors observed that the column generation restricted master problem becomes difficult to solve because it has too many constraints. Therefore, a Benders decomposition approach is proposed where the crew scheduling variables are projected out into a subproblem. The Benders decomposition approach is embedded into a branch-and-bound algorithm. The LP relaxation at each node is solved by Benders decomposition, and both the Benders master problem and the Benders subproblem are solved using column generation. The solution method is tested on instances derived from real data. The proposed solution method is compared with a pure column generation approach. The results show that the combined Benders decomposition and column generation approach produces integer solutions faster than the pure column generation method. The obtained solutions are compared with

the solutions of the traditional sequential planning process. The results show that the integrated approach produced significant savings in comparison to the traditional sequential one.

Two Benders decomposition methods are proposed and compared in Mercier et al. (2005) for the integrated crew scheduling and aircraft routing problem with restricted connection arcs. The methods are based on the combined Benders decomposition and column generation method proposed in Cordeau et al. (2001). The first decomposition considers the aircraft routing problem as the master problem and the crew scheduling problem as the subproblem, like it is done in Cordeau et al. (2001), while in the second one, the decomposition is reversed. The effect of Pareto-optimal cuts on the convergence of these two Benders decomposition approaches is analyzed. The results show that Pareto-optimal cuts accelerate the convergence of Benders decomposition. The results also show that the second decomposition outperforms the one proposed in Cordeau et al. (2001).

Shao et al. (2015) proposed a combined Benders decomposition and column generation approach for the integrated crew scheduling and aircraft routing problem with fleet assignment. The Benders decomposition approach is enhanced by several acceleration techniques. In addition, a stabilization technique is used for the column generation procedure of the crew pairing subproblem. The proposed Benders decomposition approach is tested on real-world data obtained from a U.S.-based airline carrier. The results show that the integrated approach yields 8.4% improvement, on average, in comparison with a traditional sequential decision process.

Cacchiani and Salazar-González (2016) proposed two different exact methods for path-based and arc-based formulations of the integrated crew scheduling and aircraft routing problem. Both solution methods include three main steps: 1) solving the LP relaxation of the corresponding model to optimality using column generation on the path-based model and consequently finding a lower bound; 2) running a primal heuristic to obtain an upper bound; 3) finding the optimal solution. The third step for the path-based formulation consists of a branch-and-price algorithm, while for the arc-path-based model, a restricted mixed-integer program is used. A single bounding cut that significantly accelerates the solution process is also presented. The proposed algorithms are tested on real-world instances. The results show that the proposed method for the arc-path-based model outperforms not only the one proposed for the path-based model, but also the heuristic method proposed in Salazar-González (2014).

Zeighami and Soumis (2017) proposed a solution methodology based on Benders decomposition and column generation for the integrated crew pairing and assignment problem. The pairings are generated by the Benders master problem. The monthly schedules for pilots and copilots are generated by the Benders subproblems. Master problem and subproblems are solved by column generation.

## 8.2. Heuristic Solution Methods

For the heuristic solution methods, researchers mostly focus either on combining heuristic and mathematical programming approaches to come up with matheuristics (Crainic et al., 2014, 2018; Zhu et al., 2014; Salazar-González, 2014; Belotti et al., 2008), or on developing neighborhood-based heuristics (Capone et al., 2007). The two following subsections are dedicated to reviewing these two types of heuristic methods.

### 8.2.1. Matheuristics

Crainic et al. (2014) proposed a matheuristic solution method for the service network design problem with resource management. Their proposed solution method can be described in two

main phases. The first phase is to solve the LP relaxation of the original problem, and the second one is an iterative *slope scaling* procedure. The idea of slope scaling is to iteratively solve a linear approximation of the formulation, and to use the resulting flow distribution to adjust the fixed cost approximation at the next iteration. When the slope scaling stalls, a perturbation procedure changes the initial linearization factors to start a new phase of the slope scaling procedure (see, e.g., Kim and Pardalos (1999); Crainic et al. (2004)). In the proposed solution method, the LP relaxation is solved by a column generation approach that generates cycles dynamically. The linearization factors of the slope scaling procedure are then initialized using the information from the LP relaxation phase. On a set of small-size instances, the results show that the algorithm produces high-quality solutions in comparison with a commercial MIP solver, as it only generates a small fraction of the possible cycles. The algorithm is also benchmarked on a set of large-scale instances against a column generation-based heuristic developed by the authors. The results show that the algorithm again can produce high-quality solutions.

Crainic et al. (2018) proposed a matheuristic that extends solution techniques proposed in Crainic et al. (2014). The matheuristic applies column generation to determine the set of resource cycles. To produce feasible solutions, the solution method uses both slope scaling and different matheuristics such as Archetti et al. (2008); Vu et al. (2013); De Franceschi et al. (2006); Hewitt et al. (2010). The results show that the proposed approach outperforms both a commercial MIP solver and the heuristic method in Crainic et al. (2014).

Zhu et al. (2014) proposed a matheuristic solution method based on slope scaling for the integrated rail freight service network design problem. Two approaches are proposed: a basic approach that linearizes all design variables and a dynamic approach, where the service design variables are linearized and a metaheuristic is used to generate the blocks dynamically. The approaches are tested on instances based on the setting of the main-line network of a major North American railroad. The results show that the optimality gap obtained by CPLEX increased dramatically with instance size, and that CPLEX is unable to find an optimal solution within 10 hours of CPU time. The basic approach obtained better solutions than CPLEX for more than 90% of the instances. The dynamic approach also outperformed CPLEX, but in comparison with the basic one, it performed somewhat worse on small and medium size instances due to the additional effort required by the block generation feature. However, it starts outperforming the basic approach when the instance dimensions grow.

Salazar-González (2014) proposed a two-phase matheuristic for the integrated crew scheduling and aircraft routing problem. The first phase is a greedy search to find the pairings to cover all the flights, and the second phase creates aircraft routes.

Belotti et al. (2008) proposed a Lagrangian relaxation method for a multilayer network design problem in telecommunications. A Lagrangian relaxation is used to relax the virtual flow capacity constraints. In this way, there is no relation between the flow variables and the design variables in the relaxed problem. Therefore, the relaxed problem is decomposed into shortest path subproblems for each commodity, plus one capacity assignment subproblem to determine the design variables. Since the number of virtual arcs increases exponentially when increasing the problem size, a column generation approach is used to solve the capacity assignment subproblem. A subgradient method is applied to find the Lagrangian lower bound. To find feasible solutions for the problem, a local search heuristic is developed. It starts with an initial solution and tries to improve it by rerouting the commodities on the virtual links and rerouting the virtual capacities on the physical links. The proposed Lagrangian relaxation method is able to find both lower and upper bounds for almost all the tested instances in a reasonable time.

### 8.2.2. Neighborhood-Based Heuristics

Capone et al. (2007) proposed a heuristic for a two-layer telecommunications network design problem with node capacity and multicast traffic. The heuristic constructs an initial solution using a greedy-based procedure. Then, it improves the solution using two different neighbourhood structures including 1) changing the routing of the commodities in the virtual layer, and 2) re-routing the virtual links on the physical links. The results show that the algorithm performs better with the second neighbourhood structure.

## 9. Conclusions

We proposed a taxonomy of multilayer network design problems using different features including the number of layers as well as the type and degree of connectivity requirements between layers. We also presented a state-of-the-art review of multilayer network design models and methods found in telecommunications and transportation applications. The review shows that there is one research contribution in the literature that considers one-to-many connectivity Zeighami and Soumis (2017). The review also shows that almost all contributions in the literature considered two-layer network design problems with design or flow connectivity. The only contribution that considers both flow and design connectivity requirements is Zhu et al. (2014). Studying new problems with flow-design connectivity and with more than two layers is a fascinating research direction.

From the solution methodology point of view, a first interesting research avenue is to adapt to multilayer network design the advanced solution methods proposed for single layer network design problems. In particular, Lagrangian relaxation methods have been used to solve single layer network design problems (see Holmberg and Yuan (2000), Sellmann et al. (2002), Crainic et al. (2004), and Kliwer and Timajev (2005) for example), but, as we have seen, they have been very rarely applied to address multilayer network design problems (the only exception is the work of Belotti et al. (2008)). Another interesting research avenue is to take advantage of the multilayer network design structure to derive new valid inequalities to be used in cutting plane methods. Finally, because of their inherent difficulty, multilayer network design problems can be solved by exact solution methods only for relatively small instances. Solving large-scale instances requires a combination of decomposition methods (cutting planes, Benders decomposition, column generation, Lagrangian relaxation) and metaheuristics. The development of matheuristics capable of solving multilayer network design problems of increasing complexity constitutes a major research challenge.

## Acknowledgements

While working on this project, the second author was Adjunct Professor with the Département d'informatique et de recherche opérationnelle, Université de Montréal. The first author gratefully acknowledges the support of the Université de Montréal, through its end of doctoral studies scholarship and the excellence scholarship of the Département d'informatique et de recherche opérationnelle, as well as the support of CIRRELT through its doctoral excellence scholarship. Partial funding for this project has also been provided by the Natural Sciences and Engineering Council of Canada (NSERC), through its Discovery Grant program. We also gratefully acknowledge the support of the Fonds de recherche du Québec through their infrastructure grants.

## References

- Andersen, J., Crainic, T. G., Christiansen, M., 2009a. Service network design with asset management: Formulations and comparative analyses. *Transportation Research Part C: Emerging Technologies* 17 (2), 197–207.
- Andersen, J., Crainic, T. G., Christiansen, M., 2009b. Service network design with management and coordination of multiple fleets. *European Journal of Operational Research* 193 (2), 377–389.
- Archetti, C., Speranza, M. G., Savelsbergh, M. W., 2008. An optimization-based heuristic for the split delivery vehicle routing problem. *Transportation Science* 42 (1), 22–31.
- Balakrishnan, A., Magnanti, T. L., Mirchandani, P., 1994. A dual-based algorithm for multi-level network design. *Management Science* 40 (5), 567–581.
- Balakrishnan, A., Magnanti, T. L., Shulman, A., Wong, R. T., 1991. Models for planning capacity expansion in local access telecommunication networks. *Annals of Operations Research* 33 (4), 237–284.
- Belotti, P., Capone, A., Carello, G., Malucelli, F., 2008. Multi-layer MPLS network design: The impact of statistical multiplexing. *Computer Networks* 52 (6), 1291–1307.
- Cacchiani, V., Salazar-González, J.-J., 2016. Optimal solutions to a real-world integrated airline scheduling problem. *Transportation Science* 51 (1), 250–268.
- Capone, A., Carello, G., Matera, R., 2007. Multi-layer network design with multicast traffic and statistical multiplexing. In: *IEEE GLOBECOM 2007-IEEE Global Telecommunications Conference*. IEEE Operations Center, pp. 2565–2570.
- Cohn, A. M., Barnhart, C., 2003. Improving crew scheduling by incorporating key maintenance routing decisions. *Operations Research* 51 (3), 387–396.
- Cordeau, J.-F., Pasin, F., Solomon, M. M., 2006. An integrated model for logistics network design. *Annals of Operations Research* 144 (1), 59–82.
- Cordeau, J.-F., Stojković, G., Soumis, F., Desrosiers, J., 2001. Benders decomposition for simultaneous aircraft routing and crew scheduling. *Transportation Science* 35 (4), 375–388.
- Costa, A. M., França, P. M., Lyra Filho, C., 2011. Two-level network design with intermediate facilities: An application to electrical distribution systems. *Omega* 39 (1), 3–13.
- Crainic, T. G., 2000. Service network design in freight transportation. *European Journal of Operational Research* 122 (2), 272–288.
- Crainic, T. G., Gendron, B., Hernu, G., 2004. A slope scaling/lagrangian perturbation heuristic with long-term memory for multicommodity capacitated fixed-charge network design. *Journal of Heuristics* 10 (5), 525–545.
- Crainic, T. G., Hewitt, M., Toulouse, M., Vu, D. M., 2014. Service network design with resource constraints. *Transportation Science* 50 (4), 1380–1393.
- Crainic, T. G., Hewitt, M., Toulouse, M., Vu, D. M., 2018. Scheduled service network design with resource acquisition and management. *EURO Journal on Transportation and Logistics* 7 (3), 277–309.
- Crainic, T. G., Ricciardi, N., Storch, G., 2009. Models for evaluating and planning city logistics systems. *Transportation Science* 43 (4), 432–454.
- Dahl, G., Martin, A., Stoeer, M., 1999. Routing through virtual paths in layered telecommunication networks. *Operations Research* 47 (5), 693–702.
- De Franceschi, R., Fischetti, M., Toth, P., 2006. A new ilp-based refinement heuristic for vehicle routing problems. *Mathematical Programming* 105 (2-3), 471–499.
- Fortz, B., Poss, M., 2009. An improved benders decomposition applied to a multi-layer network design problem. *Operations Research Letters* 37 (5), 359–364.
- Hewitt, M., Nemhauser, G. L., Savelsbergh, M. W., 2010. Combining exact and heuristic approaches for the capacitated fixed-charge network flow problem. *INFORMS Journal on Computing* 22 (2), 314–325.
- Holmberg, K., Yuan, D., 2000. A Lagrangian heuristic based branch-and-bound approach for the capacitated network design problem. *Operations Research* 48 (3), 461–481.
- Kim, D., Pardalos, P. M., 1999. A solution approach to the fixed charge network flow problem using a dynamic slope scaling procedure. *Operations Research Letters* 24 (4), 195–203.
- Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J. P., Moreno, Y., Porter, M. A., 2014. Multilayer networks. *Journal of Complex Networks* 2 (3), 203–271.
- Kliwer, G., Timajev, L., 2005. Relax-and-cut for capacitated network design. In: Støtting Brodal, G., Leonardi, S. (Eds.), *Algorithms—ESA 2005, 13th Annual European Symposium*. Springer, pp. 47–58.
- Knippel, A., Lardeux, B., 2007. The multi-layered network design problem. *European Journal of Operational Research* 183 (1), 87–99.
- Koster, A. M., Orlowski, S., Raack, C., Baier, G., Engel, T., 2008. Single-layer cuts for multi-layer network design problems. In: S. Raghavan, Bruce Golden, E. W. (Ed.), *Telecommunications Modeling, Policy, and Technology*. Springer Science and Business Media, Ch. 1, pp. 1–23.

- Lin, C.-C., 2010. The integrated secondary route network design model in the hierarchical hub-and-spoke network for dual express services. *International Journal of Production Economics* 123 (1), 20–30.
- Magnanti, T. L., Wong, R. T., 1984. Network design and transportation planning: Models and algorithms. *Transportation Science* 18 (1), 1–55.
- Mattia, S., 2012. Solving survivable two-layer network design problems by metric inequalities. *Computational Optimization and Applications* 51 (2), 809–834.
- Mattia, S., 2013. A polyhedral study of the capacity formulation of the multilayer network design problem. *Networks* 62 (1), 17–26.
- Mercier, A., Cordeau, J.-F., Soumis, F., 2005. A computational study of benders decomposition for the integrated aircraft routing and crew scheduling problem. *Computers & Operations Research* 32 (6), 1451–1476.
- Mercier, A., Soumis, F., 2007. An integrated aircraft routing, crew scheduling and flight retiming model. *Computers & Operations Research* 34 (8), 2251–2265.
- Minoux, M., 1989. Networks synthesis and optimum network design problems: Models, solution methods and applications. *Networks* 19 (3), 313–360.
- Obreque, C., Donoso, M., Gutiérrez, G., Marianov, V., 2010. A branch and cut algorithm for the hierarchical network design problem. *European Journal of Operational Research* 200 (1), 28–35.
- Orlowski, S., 2009. Optimal design of survivable multi-layer telecommunication networks. Ph.D. thesis, PhD thesis, Technische Universität Berlin.
- Orlowski, S., Raack, C., Koster, A. M., Baier, G., Engel, T., Belotti, P., 2010. Branch-and-cut techniques for solving realistic two-layer network design problems. In: Arie Koster, X. M. (Ed.), *Graphs and Algorithms in Communication Networks*. Springer Berlin Heidelberg, Ch. 3, pp. 95–118.
- Orlowski, S., Wessälly, R., 2004. An integer programming model for multi-layer network design. Tech. Rep. 04-49, ZIB, Berlin.
- Raack, C., Koster, A. M., 2009. An integer set-packing problem arising in two-layer network design. In: *Proceedings of INOC 2009*.
- Salazar-González, J.-J., 2014. Approaches to solve the fleet-assignment, aircraft-routing, crew-pairing and crew-rostering problems of a regional carrier. *Omega* 43, 71–82.
- Sellmann, M., Kliewe, G., et al., 2002. Lagrangian cardinality cuts and variable fixing for capacitated network design. In: *Algorithms—ESA 2002*. Springer, pp. 845–858.
- Shao, S., Sherali, H. D., Haouari, M., 2015. A novel model and decomposition approach for the integrated airline fleet assignment, aircraft routing, and crew pairing problem. *Transportation Science* 51 (1), 233–249.
- Vu, D. M., Crainic, T. G., Toulouse, M., 2013. A three-phase matheuristic for capacitated multi-commodity fixed-cost network design with design-balance constraints. *Journal of heuristics* 19 (5), 757–795.
- Zeighami, V., Soumis, F., 2017. Combining Benders decomposition and column generation for integrated crew pairing and personalized crew assignment problems. Tech. Rep. G–2017–41, Cahiers du GERAD.
- Zhu, E., Crainic, T. G., Gendreau, M., 2014. Scheduled service network design for freight rail transportation. *Operations Research* 62 (2), 383 – 400.