

Multi-period Hub Network Design Problems with Modular Capacities

Sibel A. Alumur^{a,b}, Stefan Nickel^{c,d}, Francisco Saldanha-da-Gama^e, Yusuf Seçer^{b,f}

sibel.alumur@uwaterloo.ca, stefan.nickel@kit.edu, fsgama@fc.ul.pt, yusufsecerdin@gmail.com

^a Department of Management Sciences, University of Waterloo, Waterloo, Ontario, Canada

^b Department of Industrial Engineering, TOBB University of Economics and Technology, Ankara, Turkey

^c Institute for Operations Research, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

^d Fraunhofer Institute for Industrial Mathematics (ITWM), Kaiserslautern, Germany

^e DEIO-CIO, Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal

^f Department of Industrial Engineering, University of Miami, Coral Gables, Florida, USA

Abstract

In this paper, a modeling framework is proposed for multi-period hub location. The problems to be addressed are extensions of classical hub location problems to the situation in which the hub network can be progressively built and its capacity gradually expanded over time. Both the single allocation and the multiple allocation cases are considered. For each case, a mixed-integer linear programming formulation is proposed and a set of valid inequalities is derived for enhancing the corresponding model. The results of a set of computational tests performed using the formulations proposed and their enhancements are reported. The value of the multi-period solution is discussed as a measure for evaluating the relevance of considering a multi-period model instead of a static counterpart.

1 Introduction

The basic setting underlying a hub location problem consists in a set of nodes such that some commodity or flow must be shipped between each origin-destination (O-D) pair. One possibility would be to make direct shipments between all O-D pairs. However, not only this is often not implementable due to limited resources but also significant savings in costs as well as gains in efficiency can be achieved by considering the alternative which consists in selecting a subset of nodes to become hubs, i.e., a subset of nodes that consolidate and redistribute the flow in the network. The goal is to find the network design which minimizes the total setup cost for the hubs plus the total flow routing cost.

One important aspect in this type of network design problems concerns the allocation of the non-hub nodes to the hubs. Two allocation patterns have been mostly considered in the literature (the reader should refer to Alumur and Kara [1] for further details): single allocation and multiple allocation. In the former case, each non-hub node is allocated to exactly one hub; in the latter, a non-hub node can be allocated to more than one hub.

Another relevant feature of a hub location problem regards the subnetwork induced by the hubs. Initially, authors often considered that such subnetwork was a complete graph. This is the case, for instance, in the works by Campbell [7], Ernst and Krishnamoorthy [15] and Ernst and Krishnamoorthy [17]. However, more recently, due to practical needs, such structure has been relaxed and incomplete hub networks have been considered in the literature. In this case, the design of the hub network is also part of the decision making process. Some works in this direction include Alumur et al. [2], Alumur et al. [3], Contreras et al. [11], Gelareh and Nickel [19], Nickel et al. [22], Yaman [25] and Yoon and Current [26].

Hub location problems have great practical relevance with applications arising in areas such as postal delivery services (Ernst and Krishnamoorthy [16]), public transportation networks (Nickel et al. [22]) and logistics systems (Tan and Kara [24]). Since the seminal work by O’Kelly [23] much research has been done on this type of problems. The reader can refer to the reviews by Campbell et al. [8], Alumur and Kara [1], and Campbell and O’Kelly [9] for an overview on the area.

A hub location problem is often solved as part of a strategic decision making process and its solution in terms of the selected hubs is often expected to last for a considerable amount of time. During this time, the data underlying the problem may change. This is the case, for instance, with the flows/demands. Furthermore, building a hub network is often a time consuming activity which may prevent it from being accomplished in a single step. In this situation, the network structure must be progressively built over time. Finally, the combination of the previous aspects may require an adjustment of the operating capacity over time. As a consequence, locating a set of hubs becomes a question not only of “where” but also of “when”. This reveals the multi-period nature of this type of problems, which has often been disregarded in the literature. This additional feature is the main focus of this paper. The multi-period nature of many practical facility location problems was particularly emphasized by Melo et al. [20]. Hub location problems are no exception and this work aims at fulfilling the gap existing in the area.

The literature on multi-period hub location problems is scarce. To the best of the authors’ knowledge only two papers exist addressing hub location problems featuring this aspect.

Campbell [7] proposes a continuous approximation model for a dynamic hub location problem. Demand increases over time. A discrete set of time periods is considered in which new hubs can be located and existing hubs relocated. A set of origin and destination points is considered which is assumed to be scattered over the service region. The service region is assumed to be finite and of fixed size and shape. The performance measure to be optimized is the present value of the sum of the terminal, transportation and relocation costs, discounted over an infinite time horizon. Several myopic strategies for locating the hubs (strategies which are based on the current demands and terminal locations) are investigated. Their performance is evaluated using a lower bound on the optimal cost, which, in turn, is obtained ignoring the relocation costs. The work focuses on single and double dimensional spaces. Hubs are uncapacitated and multiple allocation is allowed.

Contreras et al. [10] address a discrete multiple allocation hub location problem in a multi-period setting. The goal is to allow adjustments of the hub network over time. A planning horizon is considered which is partitioned into a finite set of time periods. In each time period, a decision has to be made regarding the nodes that should become hubs. No relation is assumed for the set of operating hubs in consecutive periods. For instance, a node can be a hub in non-consecutive periods. Flows are assumed to be known for each period of the planning horizon. Hubs are uncapacitated. A branch-and-bound procedure is proposed for solving the problem. At each node of the tree, lower and upper bounds are obtained using lagrangean relaxation.

In addition to the references above, we can also cite the PhD thesis by Gelareh [18] where a model was proposed for a phase-in/phase-out multi-period hub location problem. In each time period, new hubs can be built and hubs that existed at the beginning of the planning horizon can be removed. The set of hub links that should operate in each period is also part of the decision making process. Due to the complexity of the problem and to the fact that only toy instances can be solved to optimality using a general solver, a local search based approach is developed for obtaining feasible solutions to the problem.

In this work, we consider the basic setting underlying a hub location problem, i.e., a set of nodes such that some flow has to be shipped between each O-D pair. We consider a planning horizon partitioned into a set of time periods (e.g. years or quarters of a year) and we assume that the flows change over time but that they have previously been forecasted leading to a flow matrix for each time period. Like in the literature on static hub location problems, we consider a pure phase-in problem that is, we assume that a new hub network should be built from scratch. However, in this work, this is done progressively over time. Nevertheless, as it will be discussed later on, the modeling framework we propose can be easily adapted to the situation in which

a network is already operating and its expansion (in terms of the hubs and of the operating capacity of the exiting hubs) is to be planned.

We assume that hubs are capacitated and that capacities are modular. The reason for assuming modular capacities has to do with the practical applications of hub location problems in which capacities are often discrete in the sense that a facility is never expanded continuously but by the addition of modules (e.g. a new terminal in an airport or a new sorting line in a postal delivery system).

Finally, in terms of the capacity, we assume that for each hub a limit exists on the total flow that goes through the hub. This includes the flow originated at the hub. It is important to note that when hubs are capacitated it is up to the type of resources or operations involved in the problem to determine the exact shape of the capacity constraints. The reader should refer to Correia et al. [13] for a deeper discussion on this issue. In our case, we assume that a hub must collect its own out-bound flow and distribute its own in-bound flow and thus, capacity refers to all this flow.

In this work, we address both the single allocation and the multiple allocation patterns. Such patterns are period-dependent, i.e., the allocation of a non-hub node to a hub (in the case of single allocation) or to multiple hubs (in the case of multiple allocation) can change over time.

Table 1 emphasizes the distinguishing features between this study and the other multi-period hub location studies that can be found in the literature namely those by Campbell [7], Contreras et al. [10], and Gelareh [18]. As the last column of Table 1 reveals, the current paper is the first one considering hub capacities in a multi-period setting. We additionally incur hub network design decisions and do not assume a fully interconnected hub network. Moreover, we consider an operational cost per unit of flow at a hub facility, which, again, is done for the first time within a multi-period context.

Table 1: Comparison of multi-period hub location studies.

Problem Features	Campbell [7]	Contreras et al. [10]	Gelareh [18]	This study
Location space	Continuous	Discrete	Discrete	Discrete
Allocation rule	Single & Multiple	Multiple	Multiple	Single & Multiple
Capacitated hubs	—	—	—	✓
Hub network design	—	—	✓	✓
Cost parameters				
Setup cost for hubs	✓	✓	✓	✓
Transportation cost	✓	✓	✓	✓
Cost of operating a hub link	—	—	✓	✓
Operational cost at the hubs	—	—	—	✓

For the single and multiple allocation problems considered, we propose mixed-integer linear

programming formulations which result from the extension to the multi-period setting of well-known models that have been developed for the static case. Additionally, several sets of valid inequalities are derived in order to enhance the developed models. We report the results of a set of computational tests performed using the models proposed and their enhancements. Finally, we discuss the value of the multi-period solution, which is a measure that we propose for evaluating the relevance of using a multi-period model instead of a static counterpart.

The remainder of the paper is organized as follows. In the next section, we introduce a modeling framework for multi-period hub location and introduce several sets of valid inequalities which can be used for enhancing the proposed models. In section 3, we report and discuss the results of the computational tests performed using the models proposed as well as the valid inequalities derived. Finally, in section 4 we discuss the value of the multi-period solution. The paper ends with some conclusions drawn from the work done.

2 Mathematical Models

The basic assumptions underlying the multi-period setting we consider for hub location are the following:

- The planning horizon is finite and divided into time periods.
- In each time period it is possible to open new hubs and expand the capacity of existing hubs.
- During the planning horizon it is not possible to close an existing hub or reduce the operating capacities. As mentioned before, we are planning for designing a hub network from scratch (as usually done in the literature) but we assume that it can be done progressively over time.
- Different hub links can be operated in different periods, i.e., the hub network can change over time even when the set of hubs remains unchanged.
- A hub must collect its own out-bound flow and distribute its own in-bound flow. As a consequence, all the flow routed between hubs is discounted. Note that this is not the case in some static problems addressed in the literature. For instance, in the static multiple allocation hub location problem studied by Boland et al. [6] and Ebery et al. [14] a hub does not have to collect the flow it originates (its out-bound flow). In these works, the authors consider that the capacity constraints refer to the non-processed incoming flow

(collected flow). This leads to some flow being carried out between hub nodes without any economies of scale discount (since such flow is not processed).

- Capacity constraints are applied to the volume of flow entering a hub via collection and transfer. Note that since the problem involves hub network design decisions, the flow between an origin-destination pair can be routed via more than two hubs. We assume that the flow entering a hub from any other hub may require some additional sorting. This is why the flow entering a hub via transfer is also included in capacity constraints.
- Capacities are modular. In particular, for each node, a set of modules of different sizes is assumed to be available (in case the node is selected to become a hub). In each time period, at most one module can be installed/added in each node.
- In each time period, each demand node can send flow to any other demand node through the network. Under this full cross-traffic assumption and since a hub must collect its own out-bound flow and distribute its own in-bound flow, the hub network to be established needs to be connected in each period.
- Single allocation and multiple allocation patterns are considered (separately). In both cases, it is possible to allocate a non-hub node to different hubs in different periods.
- The costs involved in the problem include i) fixed setup costs for locating the hubs, ii) fixed costs for operating the hub links, iii) fixed costs for installing modules in the hubs, iv) variable operational costs for the flow in the hubs, and v) variable costs for routing the flow between adjacent nodes.

The decisions to be made in each period regard the design of the hub network and the routing of the flow. We assume that all the decisions are made at the beginning/end of the time periods. The goal is to minimize the total costs over the planning horizon.

Before modeling the problem, we introduce the following notation:

Sets:

N	Set of nodes.
T	Set of time periods.

Flows:

w_{ij}^t	Flow originated at node $i \in N$ that is destined to node $j \in N$ in period $t \in T$.
$O_i^t = \sum_{j \in N} w_{ij}^t$	Total flow originated at node $i \in N$ in period $t \in T$.
$D_i^t = \sum_{j \in N} w_{ji}^t$	Total flow destined to node $i \in N$ in period $t \in T$.

Capacities:

Q_k	Set of different module types available for a hub located at node $k \in N$.
Γ_k^q	Capacity of module type $q \in Q_k$ available for node $k \in N$.

Costs:

f_k^t	Fixed setup cost for locating a hub at node $k \in N$ in period $t \in T$.
g_{kl}^t	Fixed cost of operating a hub link between hubs $k \in N$ and $l \in N$ in period $t \in T$.
h_k^{qt}	Cost for installing a module of type $q \in Q_k$ at hub $k \in N$ in period $t \in T$.
p_k^t	Operational cost per unit of flow for hub $k \in N$ in period $t \in T$.
c_{ij}^t	Cost of sending one unit of flow from node $i \in N$ to node $j \in N$ in period $t \in T$.
α	Economies of scale discount factor for the flow transferred between hubs.

In order to formulate the problem, we consider as a starting point well-known formulations for static capacitated single and multiple allocation hub location problems. In particular, we focus on the formulation proposed by Ernst and Krishnamoorthy [15] for the single allocation case as well as the formulation by Boland et al. [6] and Ebery et al. [14] for the multiple allocation case. Several studies suggest that these formulations represent a good trade-off between compactness and tightness, which is the reason why we consider them as a basis for our work. Note, however, that the models we present below are not straightforward extensions of these formulations because in addition to the multi-period setting a decision has to be made concerning the hub network design. Furthermore, the capacity constraints we consider require additional caution as we show below. We address the single allocation and multiple allocation cases separately in the following subsections.

2.1 The Single Allocation Problem

In order to capture the decisions that have to be made in the multi-period capacitated single allocation hub location problem we are addressing, the following sets of decision variables are proposed:

$$x_{ik}^t = \begin{cases} 1, & \text{if node } i \text{ is assigned to hub } k \text{ in period } t \\ 0, & \text{otherwise} \end{cases} \quad (i, k \in N, t \in T)$$

($x_{kk}^t = 1$ indicates that node $k \in N$ is a hub node in period $t \in T$.)

$$z_{kl}^t = \begin{cases} 1, & \text{if a hub link is operated between hubs } k \text{ and } l \text{ in period } t \\ 0, & \text{otherwise} \end{cases} \quad (k, l \in N : k < l, t \in T)$$

$$u_k^{qt} = \begin{cases} 1, & \text{if a module of type } q \text{ is installed at hub } k \text{ in period } t \\ 0, & \text{otherwise} \end{cases} \\ (k \in N, q \in Q_k, t \in T)$$

y_{kl}^{it} = Amount of flow originated at node i that is routed from hub k to hub l in period t ($i, k, l \in N, t \in T$).

The problem can be formulated as follows:

$$(P_1) \quad \text{Min} \quad \sum_{t \in T} \sum_{i \in N} \sum_{k \in N} c_{ik}^t O_i^t x_{ik}^t + \sum_{t \in T} \sum_{i \in N} \sum_{k \in N} \sum_{l \in N} \alpha c_{kl}^t y_{kl}^{it} + \sum_{t \in T} \sum_{i \in N} \sum_{k \in N} c_{ki}^t D_i^t x_{ik}^t \\ + \sum_{t \in T} \sum_{k \in N} p_k^t \left(\sum_{i \in N} O_i^t x_{ik}^t + \sum_{l \in N} \sum_{i \in N} y_{lk}^{it} \right) + \sum_{t \in T} \sum_{k \in N} \sum_{q \in Q_k} h_k^{qt} u_k^{qt} \\ + \sum_{t \in T} \sum_{k \in N} f_k^t (x_{kk}^t - x_{kk}^{t-1}) + \sum_{t \in T} \sum_{k \in N} \sum_{l \in N : k < l} g_{kl}^t z_{kl}^t \quad (1)$$

$$\text{s. t.} \quad \sum_{k \in N} x_{ik}^t = 1 \quad \forall i \in N, t \in T \quad (2)$$

$$x_{ik}^t \leq x_{kk}^t \quad \forall i, k \in N, t \in T \quad (3)$$

$$x_{kk}^{t-1} \leq x_{kk}^t \quad \forall k \in N, t \in T \quad (4)$$

$$\sum_{q \in Q_k} u_k^{qt} \leq x_{kk}^t \quad \forall k \in N, t \in T \quad (5)$$

$$\sum_{i \in N} O_i^t x_{ik}^t + \sum_{l \in N} \sum_{i \in N} y_{lk}^{it} \leq \sum_{\tau=1}^t \sum_{q \in Q_k} \Gamma_k^q u_k^{q\tau} \quad \forall k \in N, t \in T \quad (6)$$

$$\sum_{l \in N, l \neq k} y_{kl}^{it} - \sum_{l \in N, l \neq k} y_{lk}^{it} = O_i^t x_{ik}^t - \sum_{j \in N} w_{ij}^t x_{jk}^t \quad \forall i, k \in N, t \in T \quad (7)$$

$$z_{kl}^t \leq x_{kk}^t \quad \forall k, l \in N : k < l, t \in T \quad (8)$$

$$z_{kl}^t \leq x_{ll}^t \quad \forall k, l \in N : k < l, t \in T \quad (9)$$

$$y_{kl}^{it} + y_{lk}^{it} \leq O_i^t z_{kl}^t \quad \forall i, k, l \in N : k < l, t \in T \quad (10)$$

$$x_{ik}^t \in \{0, 1\} \quad \forall i, k \in N, t \in T \quad (11)$$

$$x_{kk}^0 = 0 \quad \forall k \in N \quad (12)$$

$$z_{kl}^t \in \{0, 1\} \quad \forall k, l \in N : k < l, t \in T \quad (13)$$

$$u_k^{qt} \in \{0, 1\} \quad \forall k \in N, q \in Q_k, t \in T \quad (14)$$

$$y_{kl}^{it} \geq 0 \quad \forall i, k, l \in N, t \in T \quad (15)$$

The objective function (1) accounts for the total cost over the planning horizon (transportation, operation of hubs, capacity establishment, location of hubs, and operation of hub links). Constraints (2) assure that in each period, each node is either a hub or is allocated to a hub. Constraints (3) guarantee that in each period nodes should be allocated to operating hubs. Constraints (4) impose that once installed, a hub should not be removed during the planning horizon. Inequalities (5) state that in each time period at most one module can be installed/added to each hub. Constraints (6) are the capacity constraints. Note again that the capacity of a hub refers to the incoming flow from the demand nodes (collection flow) as well as flow from other hubs (transfer). Furthermore, the capacity of a hub in each period is the sum of the capacities of all the modules established at the node until that period. Constraints (7) are flow divergence constraints. The condition that if a hub link is being operating in some period, then the corresponding end nodes should in fact be hubs is assured by constraints (8) and (9). Constraints (10) are consistency constraints between the flow variables y and the variables associated with the operating hub links z . Finally, (11)-(15) are domain constraints.

The model above can be used for the situation in which a network is already operating and the goal is to plan for its expansion. This meets many practical situations such as in public transportation systems where the expansion of an existing network may be the relevant aspect to consider. In order to address this case, in addition to the previous notation we also consider:

- N^0 Set of hubs which are operating before the beginning of the planning horizon.
- u_k^0 Operating capacity at hub $k \in N^0$ before the beginning of the planning horizon.

The new situation can now be accommodated in the model above by setting $x_{kk}^0 = 1$ for $k \in N^0$ and by replacing (6) with

$$\sum_{i \in N} O_i^t x_{ik}^t + \sum_{l \in N} \sum_{i \in N} y_{lk}^{it} \leq \sum_{\tau=1}^t \sum_{q \in Q_k} \Gamma_k^q u_k^{q\tau} \quad \forall k \in N \setminus N^0, t \in T \quad (16)$$

and

$$\sum_{i \in N} O_i^t x_{ik}^t + \sum_{l \in N} \sum_{i \in N} y_{lk}^{it} \leq u_k^0 + \sum_{\tau=1}^t \sum_{q \in Q_k} \Gamma_k^q u_k^{q\tau} \quad \forall k \in N^0, t \in T \quad (17)$$

2.2 The Multiple Allocation Problem

In this section we propose a model for the multiple allocation problem. In order to capture the decisions to be made we consider the following sets of decision variables:

$$\xi_k^t = \begin{cases} 1, & \text{if node } k \text{ is a hub node in period } t \\ 0, & \text{otherwise} \end{cases} \quad (k \in N, t \in T)$$

$$z_{kl}^t = \begin{cases} 1, & \text{if a hub link is operated between hubs } k \text{ and } l \text{ in period } t \\ 0, & \text{otherwise} \end{cases} \quad (k, l \in N : k < l, t \in T)$$

$$u_k^{qt} = \begin{cases} 1, & \text{if a module of type } q \text{ is installed at hub } k \text{ in period } t \\ 0, & \text{otherwise} \end{cases} \\ (k \in N, q \in Q_k, t \in T)$$

$$x_{ik}^t = \text{Amount of flow originated at node } i \text{ that is routed from hub } k \text{ in period } t \\ (i, k \in N, t \in T).$$

$$v_{lj}^{it} = \text{Amount of flow originated at node } i \text{ that is routed to node } j \text{ via hub } l \text{ as the} \\ \text{last hub on the route in period } t \ (i, l, j \in N, t \in T).$$

$$y_{kl}^{it} = \text{Amount of flow originated at node } i \text{ that is routed from hub } k \text{ to hub } l \text{ in} \\ \text{period } t \ (i, k, l \in N, t \in T).$$

The multi-period capacitated multiple allocation hub location problem can be formulated as follows:

$$(P_2) \quad \text{Min} \quad \sum_{t \in T} \sum_{i \in N} \sum_{k \in N} c_{ik}^t x_{ik}^t + \sum_{t \in T} \sum_{i \in N} \sum_{k \in N} \sum_{l \in N} \alpha c_{kl}^t y_{kl}^{it} + \sum_{t \in T} \sum_{i \in N} \sum_{l \in N} \sum_{j \in N} c_{lj}^t v_{lj}^{it} \\ + \sum_{t \in T} \sum_{k \in N} p_k^t \left(\sum_{i \in N} x_{ik}^t + \sum_{l \in N} \sum_{i \in N} y_{lk}^{it} \right) + \sum_{t \in T} \sum_{k \in N} \sum_{q \in Q_k} h_k^{qt} u_k^{qt} \\ + \sum_{t \in T} \sum_{k \in N} f_k^t (\xi_k^t - \xi_k^{t-1}) + \sum_{t \in T} \sum_{k \in N} \sum_{l \in N: k < l} g_{kl}^t z_{kl}^t \quad (18)$$

$$s. t. \quad \sum_{k \in N} x_{ik}^t = O_i^t \quad \forall i \in N, t \in T \quad (19)$$

$$\sum_{l \in N} v_{lj}^{it} = w_{ij}^t \quad \forall i, j \in N, t \in T \quad (20)$$

$$\sum_{q \in Q_k} u_k^{qt} \leq \xi_k^t \quad \forall k \in N, t \in T \quad (21)$$

$$\xi_k^{t-1} \leq \xi_k^t \quad \forall k \in N, t \in T \quad (22)$$

$$\sum_{i \in N} x_{ik}^t + \sum_{l \in N} \sum_{i \in N} y_{lk}^{it} \leq \sum_{\tau=1}^t \sum_{q \in Q_k} \Gamma_k^q u_k^{q\tau} \quad \forall k \in N, t \in T \quad (23)$$

$$\sum_{l \in N, l \neq k} y_{kl}^{it} - \sum_{l \in N, l \neq k} y_{lk}^{it} = x_{ik}^t - \sum_{j \in N} v_{kj}^{it} \quad \forall i, k \in N, t \in T \quad (24)$$

$$x_{ik}^t \leq O_i^t \xi_k^t \quad \forall i, k \in N, t \in T \quad (25)$$

$$x_{ik}^t \leq O_i^t (1 - \xi_i^t) \quad \forall i, k \in N : i \neq k, t \in T \quad (26)$$

$$\sum_{i \in N} v_{ij}^{it} \leq D_j^t \xi_l^t \quad \forall j, l \in N, t \in T \quad (27)$$

$$v_{lj}^{it} \leq D_j^t (1 - \xi_j^t) \quad \forall i, j, l : l \neq j, t \in T \quad (28)$$

$$z_{kl}^t \leq \xi_k^t \quad \forall k, l \in N : k < l, t \in T \quad (29)$$

$$z_{kl}^t \leq \xi_l^t \quad \forall k, l \in N : k < l, t \in T \quad (30)$$

$$y_{kl}^{it} + y_{lk}^{it} \leq O_i^t z_{kl}^t \quad \forall i, k, l \in N : k < l, t \in T \quad (31)$$

$$\xi_k^t \in \{0, 1\} \quad \forall k \in N, t \in T \quad (32)$$

$$\xi_k^0 = 0 \quad \forall k \in N \quad (33)$$

$$u_k^{qt} \in \{0, 1\} \quad \forall k \in N, q \in Q_k, t \in T \quad (34)$$

$$z_{kl}^t \in \{0, 1\} \quad \forall k, l \in N : k < l, t \in T \quad (35)$$

$$x_{ik}^t \geq 0 \quad \forall i, k \in N, t \in T \quad (36)$$

$$y_{kl}^{it} \geq 0 \quad \forall i, k, l \in N, t \in T \quad (37)$$

$$v_{lj}^{it} \geq 0 \quad \forall i, l, j \in N, t \in T \quad (38)$$

The objective function (18) accounts for the total cost over the planning period (transportation, operation of the hubs, capacity establishment, location of hubs, and operation of hub links). Constraints (19) assure that in every period of the planning horizon all the flow originated at each node should be routed via at least one hub. Constraints (20) guarantee that all flow destined to each node should be distributed via at least one hub. Constraints (21) state that in each period, at most one module can be installed/added to a hub. The assumption that we are dealing with a pure phase-in problem is assured by constraints (22). The capacity constraints are guaranteed by the set of inequalities (23). The flow conservation and divergence constraints are described by equalities (24). Constraints (25) and (26) are consistency constraints for the flow collected at the nodes. In particular, they assure that in each period, the flow originated at a demand node can be collected at a hub node only if that hub node has been previously established (constraints (25)) and that a hub must collect its own out-bound flow (constraints (26)). Constraints (27) and (28) provide the necessary consistency for the distribution decision variables. By constraints (27), the flow destined to any demand node can be distributed by a hub node only if that hub node is established (in every time period). By constraints (28), a hub needs to distribute its own in-bound flow in every time period. Constraints (29) and (30) assure that if a hub link is operating in some time period, then a hub must have been previously

installed in each end node of the link. Constraints (31) guarantee that the transfer flow must in fact be routed using operating hub links. Finally, constraints (32)-(38) are domain constraints.

To the best of our knowledge, constraints (26) and (28) were first presented by Boland et al. [6] although these authors do not make use of such inequalities because they assume that a hub does not have to collect its own out-bound flow nor distribute its own in-bound flow.

Like in the single allocation case, the model above can be used for the situation in which a network is already operating and the goal is to expand it. We consider N^0 and u_k^0 ($k \in N^0$) as defined in page 9.

The new situation can be accommodated in the model above by setting $\xi_k^0 = 1$ for $k \in N^0$ and by replacing (23) with

$$\sum_{i \in N} x_{ik}^t + \sum_{l \in N} \sum_{i \in N} y_{lik}^{it} \leq \sum_{\tau=1}^t \sum_{q \in Q_k} \Gamma_k^q u_k^{q\tau} \quad \forall k \in N \setminus N^0, t \in T \quad (39)$$

and

$$\sum_{i \in N} x_{ik}^t + \sum_{l \in N} \sum_{i \in N} y_{lik}^{it} \leq u_k^0 + \sum_{\tau=1}^t \sum_{q \in Q_k} \Gamma_k^q u_k^{q\tau} \quad \forall k \in N^0, t \in T \quad (40)$$

2.3 Valid Inequalities

A set of computational tests performed using a commercial solver for tackling models P_1 and P_2 (which will be detailed in section 3) shows that a plain use of these models may be extremely demanding from a computational point of view. Accordingly, in this section, we propose a set of valid inequalities with the goal on enhancing the models.

The first set of inequalities arises from the observation that the hub network must be connected. Accordingly, if we have, say, K operating hubs then we must have at least $K - 1$ hub links. This can be stated for the single allocation model as follows:

$$\sum_{k \in N} x_{kk}^t - 1 \leq \sum_{k \in N} \sum_{l \in N : k < l} z_{kl}^t \quad \forall t \in T \quad (41a)$$

The multiple allocation version of the above inequalities is:

$$\sum_{k \in N} \xi_k^t - 1 \leq \sum_{k \in N} \sum_{l \in N : k < l} z_{kl}^t \quad \forall t \in T \quad (41b)$$

Another set of inequalities is the following:

$$\sum_{i \in N} O_i^t \leq \sum_{\tau=1}^t \sum_{k \in N} \sum_{q \in Q_k} \Gamma_k^q u_k^{q\tau} \quad \forall t \in T \quad (42)$$

These inequalities are valid both for the single and the multiple allocation formulations. They state that in each time period, the total capacity of the operating hubs is enough to handle the total amount of flow originated in the period. Note that these inequalities are implied by inequalities (6) (or (23)) and thus are redundant for the linear relaxation of both models P_1 and P_2 . Nevertheless, they turned out to be useful as their inclusion in the model can improve the performance of a solver when solving the models.

In addition to the inequalities above, we can go further by deriving a new set of valid inequalities regarding the establishment of capacity modules. The following inequality is valid for the single-allocation problem:

$$x_{kk}^t \leq \sum_{\tau=1}^t \sum_{q \in Q_k} u_k^{q\tau} \quad \forall t \in T \quad (43a)$$

Similarly, the following inequality is valid for the multiple-allocation problem:

$$\xi_k^t \leq \sum_{\tau=1}^t \sum_{q \in Q_k} u_k^{q\tau} \quad \forall t \in T \quad (43b)$$

These inequalities state that if a hub is operating in a period, then at least one capacity module must have been installed at that hub until the end of that period.

Computational analysis with these valid inequalities using the single and multiple allocation formulations are presented in the next section.

3 Computational Tests

We test the single and the multiple allocation capacitated multi-period hub location and hub network design formulations on the CAB data set [5]. CAB data set, which is introduced by O’Kelly [23], is based on airline passenger interactions between US cities and is frequently used to test all types of hub location problems. Values of the parameters used in our computational tests with the CAB data set are summarized in Table 2.

We test instances from the CAB data set consisting of 15 and 25 nodes. We consider a 5-period planning horizon. The flows in the first period are set to those found in the OR Library [5]. As customarily done in the literature these flows are scaled so that the total flow adds up to one. We generated two different flow scenarios to test our models. In the first scenario, flows constantly increase by 5%; i.e., in order to determine the amount of flows in a period, the flows in the previous period is multiplied by 1.05. In the second scenario, the flow for each pair of demand node is determined by multiplying the previous period’s flow by a random

Table 2: Values of the parameters on the CAB data set.

Description	Parameter	Value
Sets:		
Number of Nodes	$ N $	15, 25
Number of Periods	$ T $	5
Flows:		
Flows in the first period	w_{ij}^1	OR Library [5]
Scenario I (increasing)	w_{ij}^t	Increasing with 5%
Scenario II (random)	w_{ij}^t	$\sim U[0.9w_{ij}^{t-1}, 1.2w_{ij}^{t-1}]$
Capacity modules:		
Capacity set I (tight)	Γ_k^q	0.4, 0.5, 0.6
Capacity set II (loose)	Γ_k^q	0.5, 0.75, 1
Costs: (All costs increase by 2% in each period)		
Fixed setup cost	f_k^1	500
Fixed cost of operating a hub link	g_{kl}^1	5% of the length
Cost of installing a capacity module	h_k^{q1}	$100 \times \text{Module Capacity}$
Operational cost per unit of flow	p_k^1	1
Cost of sending one unit of flow	c_{ij}^1	OR Library [5]
Economies of scale discount factor	α	0.2, 0.4, 0.6, 0.8

value from the interval $[0.9, 1.2]$. We have two different capacity sets: tight and loose. Each capacity set contains three different capacity modules with different capacities. The capacity sets and modules are the same for each candidate hub location.

Values of all the cost parameters in the beginning of the planning horizon are listed in Table 2. All the cost values are assumed to increase by the inflation rate of 2% in each period over the planning horizon. Moreover, it is assumed that the cost of a unit capacity for each capacity module decreases by 10% with an increase in the unit capacity, resulting in economies of scale in the costs of the capacity modules.

The models are solved using IBM ILOG CPLEX 12.4 software on a 64-bit Windows 7 workstation with two Intel Xeon 2.40 GHz processors and 48 GB of RAM.

Initially, we test the performance of the valid inequalities. For this analysis, we use CAB data set with 15 nodes. Each instance is tested first without any of the valid inequalities (base model), then with the inclusion of valid inequality sets (41) and (42), and lastly with the inclusion of valid inequality sets (41), (42) and (43). The CPU time comparison with the valid inequalities is summarized in Table 3.

The first five columns of Table 3 list the values of the parameters corresponding to each instance. The remaining three columns report the CPU times required in seconds to solve the corresponding models to optimality using CPLEX. A total of 32 instances is tested.

Observe from Table 3 that the lowest CPU times are obtained with different combinations of

Table 3: CPU Time comparison with the valid inequalities.

Instance number	Allocation	Flows	Capacities	α	CPU time (sec)		
					Base model	(41)+(42)	(41)+(42)+(43)
1	Single	Scenario I (increasing)	Set I	0.2	1415.85	2025.00	934.90
2				0.4	19725.80	10608.30	18354.87
3				0.6	42568.29	13410.75	13675.63
4				0.8	5067.24	8015.18	7181.97
5			Set II	0.2	435.20	421.02	308.79
6				0.4	284.86	422.73	385.06
7				0.6	120.59	86.11	105.11
8				0.8	31.61	30.02	23.54
9		Scenario II (random)	Set I	0.2	650.40	651.07	503.35
10				0.4	5012.50	4865.81	8972.04
11				0.6	12203.49	11412.36	21130.09
12				0.8	1661.93	4577.82	4433.36
13			Set II	0.2	1224.73	309.43	468.60
14				0.4	992.68	499.34	320.24
15				0.6	126.07	104.40	86.67
16				0.8	37.99	31.12	45.55
17	Multiple	Scenario I (increasing)	Set I	0.2	1181.54	1101.62	1652.94
18				0.4	1007.89	598.53	1271.81
19				0.6	1038.84	450.20	698.10
20				0.8	251.22	195.22	296.68
21			Set II	0.2	1250.38	616.92	1063.13
22				0.4	618.73	254.31	499.53
23				0.6	700.15	303.39	531.78
24				0.8	203.91	157.34	245.01
25		Scenario II (random)	Set I	0.2	1227.87	893.56	1444.41
26				0.4	809.29	527.49	1162.08
27				0.6	933.81	509.39	684.88
28				0.8	217.09	222.29	415.10
29			Set II	0.2	892.34	594.05	954.09
30				0.4	595.99	242.21	362.62
31				0.6	392.19	236.40	557.89
32				0.8	217.78	118.25	193.29
Average					3221.82	2329.90	2780.10

the valid inequalities at different instances. Taking the first instance as an example the lowest CPU time is obtained with the model including all sets of valid inequalities. For the second and third instances, the model with the sets (41)+(42) is the one performing better. In turn, the base model (without any additional inequalities) has the best performance for the fourth instance.

Among the 32 instances listed in Table 3, the model enhanced with the valid inequality sets (41) and (42) is the fastest for 22 instances. When we look into the average CPU time requirements (last row), the set (41)+(42) has the lowest value with approximately 15 minutes difference from the models which do not include the valid inequalities.

Further analysis with the valid inequalities is presented in the Appendix. In Tables 7 and 8, for each instance, in addition to the CPU time requirement, we report the gap of the LP relaxation value, gap at the root node of the branch-and-bound tree, and the number of nodes in the branch-and-bound tree reported by CPLEX. The gap at the root node is obtained after CPLEX applies its presolve and preprocessing tools to the LP relaxation.

As expected, for all the instances, the LP relaxation gap is the lowest when all sets of valid inequalities are considered. However, (41)+(42)+(43) does not necessarily result in the lowest gap at the root node or the lowest number of nodes in the branch and bound tree at all the instances. This is explained by the preprocessing operations that the solver performs to MILP models. Even though models including all sets of valid inequalities are the strongest, the average CPU time requirement is higher than that of the models including sets (41) and (42).

Next, we solve both of our models with the CAB data set containing all of the 25 nodes. The CPU times reported above for getting the optimal solutions to the 15-node instances show that the problems we are addressing are quite difficult from a computational point of view. This is not surprising as we are considering one extra dimension - the time - to problems which are well-known to be NP-hard. One important question before addressing larger instances of the problems is then how meaningful it is to solve one instance to optimality. This question is particularly important because it is a well-known fact for many problems arising in every day life that errors are associated with data collection/forecasting. For instance, in the case of static logistics network design problems, Cordeau et al. [12] refer an error margin of at least 1% which makes it meaningless solving real life problems to optimality with a solver and thus important savings in terms of computational requirement can be achieved by stopping a solver as soon as a solution with a gap say of 1% is obtained. In the case of logistics network design problems, Melo et al. [21] emphasize this same aspect for the case of multi-period problem where the errors can be expected to increase even further as the time span for the forecasts becomes larger when

compared with static problems.

In our case, we are dealing with multi-period network design problems. Therefore, the errors associated with the data can be expected to be even larger than mentioned above as we need to consider data forecasted for moments in time which can be much ahead in the future. Accordingly, although solving the 15-node instances to optimality was important for asserting some important features such as the most relevant valid inequalities to be considered in a multi-period hub location problem, in the case of the 25-node instances we decided to set a MIP tolerance of 1.50% while solving both of our models with CPLEX. Furthermore, taking into account the information gathered when solving the 15-node instances, namely the fact that the CPU time requirement is the lowest for the majority of the instances when only inequalities (41) and (42) are considered, in the tests reported below we only considered these inequalities. Additionally, we set a time limit of 24 hours for solving each instance.

The solutions that we obtained with the single and multiple allocation problems are reported in Tables 4 and 5, respectively.

Table 4: Solutions of the single allocation problem.

Instance number	Flows	Capacities	α	Total cost	Hub locations	Hub network state (C/I)	CPU time (hr)
33			0.2	6863.47	1,4,12,17	I-I-I-I-I	7.11
34		Set I	0.4	7835.33	5,11,12,17	I-I-I-I-I	24 (4.07%)
35		(tight)	0.6	8505.57	2,4,8	C-I-I-I-I	24 (4.35%)
36	Scenario I		0.8	9035.59	4,8,20	C-I-I-I-I	24 (2.50%)
37	(increasing)		0.2	6805.36	4,12,18	I-I-I-I-I	2.76
38		Set II	0.4	7550.82	4,12,18	I-I-I-I-I	2.94
39		(loose)	0.6	8321.05	2,12,21	I-I-I-I-I	2.55
40			0.8	8890.21	4,8,20	I-I-I-I-I	1.49
41			0.2	6834.56	1,4,12,17	I-I-I-I-I	8.94
42		Set I	0.4	7789.76	5,12,18,21	I-I-I-I-I	24 (4.69%)
43		(tight)	0.6	8434.24	4,8,25	C-I-I-I-I	24 (3.62%)
44	Scenario II		0.8	8954.31	4,8,20	C-I-I-I-I	24 (2.63%)
45	(random)		0.2	6736.38	4,12,17	I-I-I-I-I	1.91
46		Set II	0.4	7500.19	4,12,17	I-I-I-I-I	3.01
47		(loose)	0.6	8261.65	4,12,25	I-I-I-I-I	3.14
48			0.8	8788.67	4,8,20	I-I-I-I-I	1.29

For each instance, Tables 4 and 5 report the objective function value, hub locations at the end of the last period, the state of the hub network in each of the five periods, and the CPU time required by CPLEX to solve the models with an MIP tolerance of 1.50%. In reporting the hub network state, ‘C’ corresponds to a complete hub network, where there is a direct link between each pair of hubs, and ‘I’ corresponds to an incomplete one. The values inside parenthesis in the last column state the final gap reported by CPLEX when a 1.50% gap was not achieved within

Table 5: Solutions of the multiple allocation problem.

Instance number	Flows	Capacities	α	Total cost	Hub locations	Hub network state (C/I)	CPU time (hr)	
49	Scenario I (increasing)	Set I (tight)	0.2	6711.22	12,17,21	I-I-I-I-I	23.19	
50			0.4	7299.97	4,12,17	I-I-I-I-I	17.86	
51			0.6	7788.42	4,12,18	I-I-I-I-I	17.32	
52			0.8	8166.84	12,18,21	I-I-I-I-I	14.01	
53		Set II (loose)	0.2	6681.91	4,12,17	I-I-I-I-I	21.08	
54			0.4	7289.44	4,12,18	I-I-I-I-I	21.18	
55			0.6	7813.32	12,18,21	I-I-I-I-I	15.72	
56			0.8	8213.18	4,12,17	I-I-I-I-I	15.64	
57		Scenario II (random)	Set I (tight)	0.2	6675.23	4,12,17	I-I-I-I-I	20.78
58				0.4	7241.57	4,12,17	I-I-I-I-I	19.61
59				0.6	7732.12	4,12,17	I-I-I-I-I	17.34
60				0.8	8099.58	4,12,18	I-I-I-I-I	12.42
61			Set II (loose)	0.2	6651.94	12,18,21	I-I-I-I-I	18.06
62				0.4	7261.27	12,18,21	I-I-I-I-I	21.78
63				0.6	7738.86	4,12,18	I-I-I-I-I	18.28
64				0.8	8137.75	12,18,21	I-I-I-I-I	15.99

24 hours.

Even though the models allow opening new hubs through the planning horizon, all the hubs are installed in the first period in all the single and multiple allocation solutions. This is explained by the fact that the costs increase by 2% in each period thus, it is cheaper to establish hubs in the beginning of the planning horizon. Even though the locations of the hubs remain the same, the capacities of the hubs, the hub network, and the allocations of the demand nodes to hubs change over time. Again because of the inflation in the cost values, the capacity modules for the hubs are established within the first two periods of the planning horizon at all the instances. Despite this tendency for pushing some important strategic decisions towards the beginning of the planning horizon, we will see in the next section that the multi-period modeling framework we propose is still preferred to a static one.

In both problems, hub locations tend to change with different values of α , different capacity sets, demand scenarios, and allocation rules. Thus, we cannot draw general conclusions on the effect of different parameter values on the resulting hub locations. However, we can say that no matter what the values of the parameters are some high demand nodes in the CAB data set, such as Chicago (4), Los Angeles (12), and New York (17), are usually preferred as hub locations.

As can be observed from the last columns of the Tables 4 and 5, the resulting hub networks at the end of the planning horizon are always incomplete. This shows that there is no need to establish a complete hub network when the operational costs of the hub links are considered.

Instances with tight capacity set (Set I) are computationally more demanding than their loose capacity (Set II) counterparts especially for the single allocation problems. The models are solved on average within 15 hours which is reasonable for a long-term strategic problem. In 6 out of the 32 instances listed in Table 4, a 1.5% gap could not be obtained within the time limit. The average final gap reported by CPLEX for these instances is 3.64% with a maximum of 4.69%.

Figure 1 depicts the resulting networks at the end of the first and the last periods for instances 33 and 36 from the single allocation problem. Similarly, Figure 2 depicts solutions of the multiple allocation problem at the end of the first and the last periods for instances 51 and 59. In both figures, the size of the squares represents the capacities of the established hubs; i.e., the bigger the hub capacity the bigger the square size.

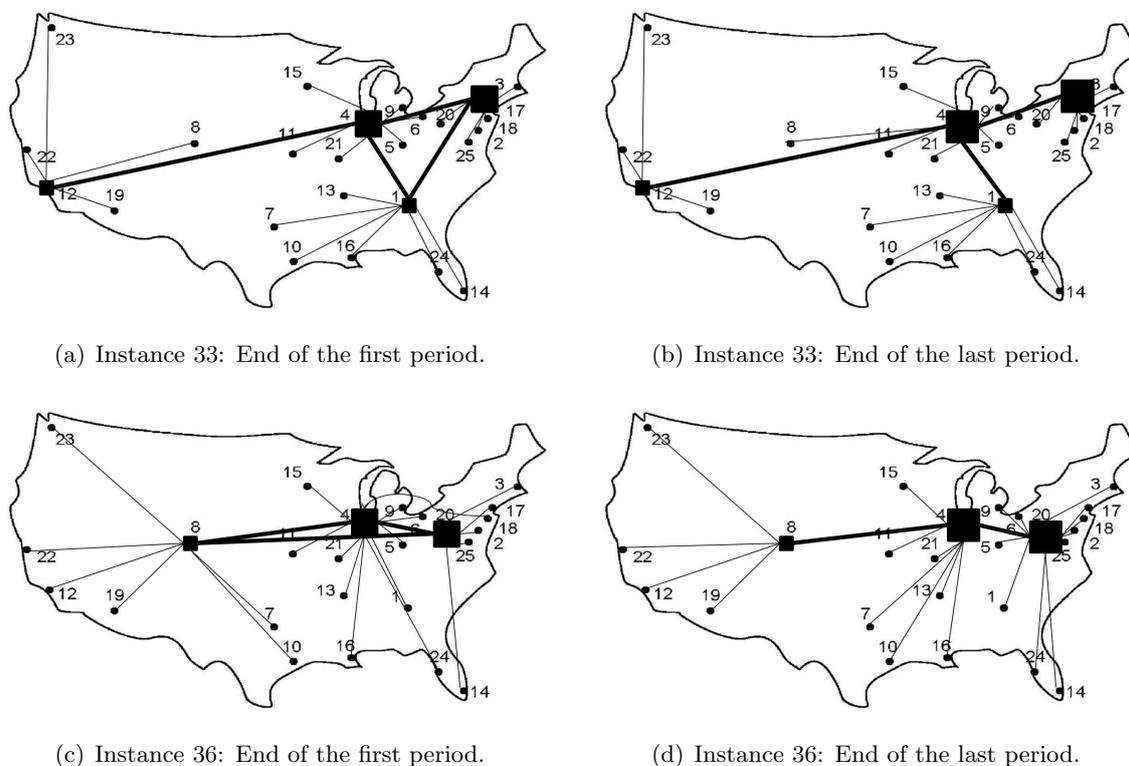


Figure 1: Some solutions for the single allocation problem.

As depicted in Figures 1(a) and 1(b) for instance 33, hubs are established at locations 1, 4, 12, and 17. An incomplete hub network is built in all periods. Four hub links are operating during the first period, whereas this number is reduced to three in the last period. Additional capacity modules are established at hubs 4 and 17 until the end of the planning horizon. With the establishment of new capacity modules, the allocation of demand node 8 changes. Node 8 is allocated to the hub located at node 12 at the end of the first period, while it is allocated to

the hub located at node 4 at the end of the last period.

Concerning instance 36 (Figures 1(c) and 1(d)), hubs are located at nodes 4, 8, and 20. The resulting hub network at the end of the first period is complete, whereas it is incomplete at the subsequent periods. Additional capacity modules are established for hubs 4 and 20. At the end of the last period, eight demand nodes (1, 5, 6, 7, 9, 10, 18, and 24) are allocated to different hub nodes when compared with the allocations at the end of the first period.

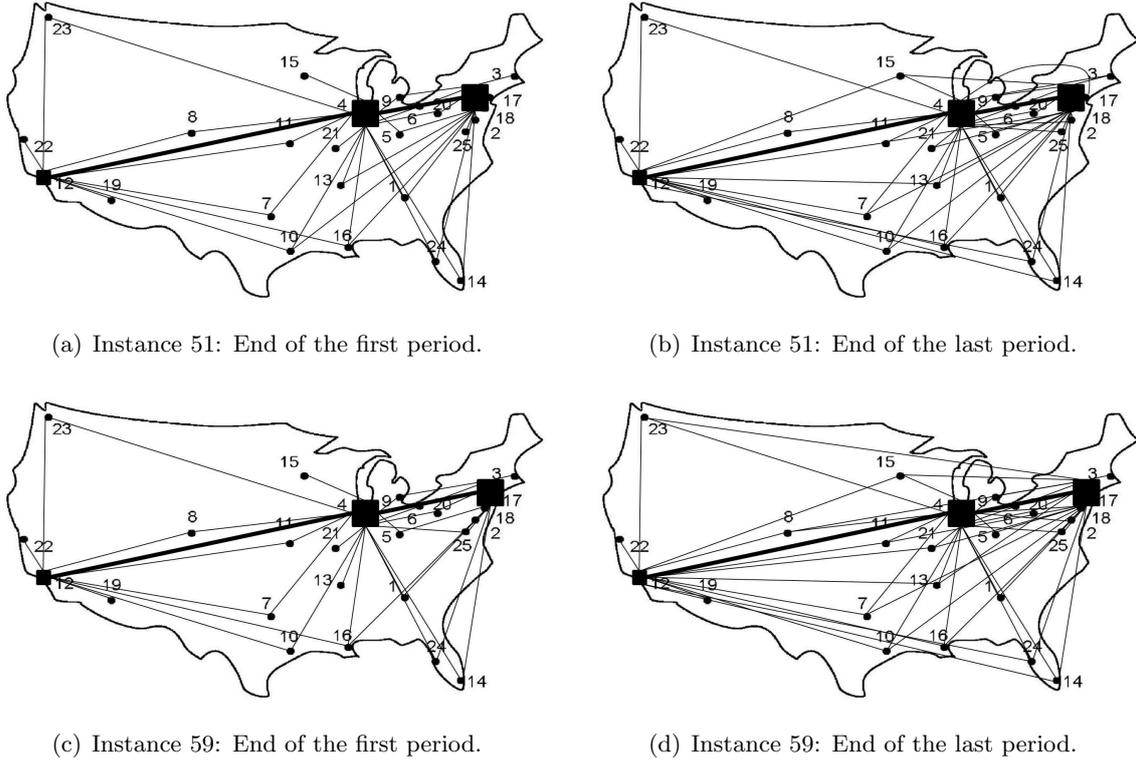


Figure 2: Some solutions for the multiple allocation problem.

In both of the multiple allocation instances depicted in Figure 2, hubs are located at nodes 4, 12, and 17. Capacity establishment decisions along the planning horizon are the same for each of these three hubs. For instance 51, at the end of the planning horizon (Figure 2(b)) nine demand nodes are allocated to additional hub nodes when compared with the allocations at the end of the first period (Figure 2(a)). Similarly, for instance 59, twelve nodes are allocated to additional hub nodes at the end of the last period (Figure 2(d)). For both instances 51 and 59, node 14 is allocated to hubs 4 and 20 at the end of the first period whereas it is allocated to all of the three hubs at the end of the last period. These instances exemplify not only the evolution of hub networks in time but also the flexibility that can be achieved by considering a multi-period modeling framework.

For some instances listed in Tables 4 and 5, neither the locations, the capacities of the hubs

nor the allocation decisions and the hub network change over time. What is more, in all of the reported results, we observe the hubs being installed in period 1. This is not surprising for two reasons: first, the setup costs for the hubs increase with time and thus, when it is necessary to use a hub during the planning horizon, it is always less costly to establish it in period 1. Second, we are considering capacities that can be adjusted over time. As a result, we observe a common situation in practice, which is the establishment of the infrastructure early in the planning horizon and then the adjustment of the capacity when necessary.

In order to check whether the models can react more drastically to important changes during the planning horizon, namely by installing hubs in other periods apart from period 1, we extended the computational tests. We considered significant changes over time in the setup costs for the hubs and left the values for all the other parameters unchanged. In particular, we considered a reduction by a factor of 0.5 for the setup costs from one period to the following starting in period 1. Such type of decrease can be observed in practice when, for instance, some significant technological change occurs or some government incentive can be assured. For this analysis, we considered two 15-node instances.

In the first case, we took the instance associated with $\alpha = 2$ for the single allocation problem using the random flow scenario and loose capacity set. The results show:

Period 1 – a single hub is installed at node 4;

Period 2 – two additional hubs are installed at nodes 7 and 12;

Period 3 – two additional hubs are installed at nodes 6 and 14;

Period 4 – one additional hub is installed at node 1.

The six hubs installed until period 4 will then operate until the end of the planning horizon.

In the second case, we considered the multiple allocation problem with $\alpha = 0.8$, increasing flow scenario and tight capacity set. The results led to:

Period 1 – two hubs are installed at nodes 4 and 7;

Period 3 – one additional hub is installed at node 6;

Once installed, the hubs operate in the remaining periods of the planning horizon.

These results show that, in fact, important changes occurring during the planning horizon may call for more drastic reactions, leading to other types of solutions.

4 The Value of the Multi-period Solution

As it was already emphasized more than once, the multi-period modeling framework proposed in Section 2 corresponds to one additional decision dimension in hub location: the time. The models tend to be large and difficult to solve even for instances with a moderate number of nodes, as the results presented in the previous section show. Accordingly, one may ask whether it is worth considering this extra dimension. In other words, consider a situation in which it is possible to build the network in a single step/period. Is it still worth considering a multi-period modeling framework when flows and costs change over time? An answer to this question can be given by making use of the *value of the multi-period solution* which is a concept first introduced by Alumur et al. [4] in the context of a multi-period reverse logistics network design problem.

The value of the multi-period solution compares the optimal value of the multi-period model with the value of a feasible solution to the multi-period problem obtained by solving a static (single-period) counterpart. A static counterpart is a problem which takes into account the information available for the planning horizon and looks for a static solution. Considering this solution, we can return to the original multi-period problem and implement it in all periods of the planning horizon. If such a solution is feasible for the multi-period problem, the difference between its value (as feasible solution to the multi-period problem) and the optimal value of the multi-period problem gives the value of the multi-period solution. In order to remove the order of magnitude of the values involved in this computation, we can consider the relative value of the multi-period solution. Denote by $v(P)$ the optimal value to the multi-period problem and denote by $v_s(P)$ the (multi-period) objective function value for the solution obtained by solving a static counterpart. The relative value of the multi-period solution is given by

$$100 \times \frac{v_s(P) - v(P)}{v(P)}$$

In general, several static counterparts can be associated with a multi-period problem. We detail this aspect in the context of multi-period hub location and define two static counterparts of our problem.

In multi-period hub location, both the flows and the costs are assumed to change over the planning horizon. The goal of a static counterpart is to define a single-period problem which finds a static solution that can be implemented at the beginning of the planning horizon remaining unchanged until the end.

One possibility for building a static counterpart is to reduce the time-dependent demand matrix to a single demand matrix by somehow condensing the information available for the different time periods. As we are facing capacitated problems, it is important to account for the

feasibility of the (static) solution in every period of the planning horizon. Accordingly, we have no alternative but to plan for the worst case, i.e., for each flow, we must consider the largest value observed over all the planning horizon. Regarding the time-dependent costs and in order to act consistently with the above mentioned procedure for the demand, we do the following: for each particular time-dependent cost component we take the largest value observed all over the planning horizon.

Once the information is condensed, we can solve the static (single-period) problem resulting from assuming that $|T| = 1$ in the models presented in Section 2. Finally, we compute the objective function value of this static solution in the multi-period problem using the parameter values for the whole planning horizon but fixing the hub location, hub capacity, hub network design and allocation decisions according to the static solution and finally, we compute the relative value of the multi-period solution as explained above. The relative difference between (i) the objective function value of the multi-period problem considering the static solution obtained by aggregating the data as described above and (ii) the optimal value of the multi-period problem is defined as the *weak value of the multi-period solution*.

Another possibility for building a static counterpart is to somehow avoid aggregating the data and use all the information on flows and costs through the whole planning horizon. To do so, we can consider additional constraints to the multi-period models stating that the network structure should remain the same in all periods of the planning horizon. In particular, we add constraints to the base models ensuring the hub and hub link establishment decisions to remain fixed during the whole planning horizon. In the case of the single-allocation problem, these constraints are the following:

$$x_{kk}^1 = x_{kk}^2 = \dots = x_{kk}^{|T|}, \quad \forall k \in N \quad (44)$$

$$z_{kl}^1 = z_{kl}^2 = \dots = z_{kl}^{|T|}, \quad \forall k, l \in N : k < l \quad (45)$$

$$u_k^{qt} = 0, \quad \forall k \in N, q \in Q_k, t = 2, \dots, |T| \quad (46)$$

In the multiple allocation case the constraints to be added are (45) and (46) as well as:

$$\xi_k^1 = \xi_k^2 = \dots = \xi_k^{|T|}, \quad \forall k \in N \quad (47)$$

This way, we end up with a (static) network structure which makes use of all available data. Now, we call the *strong value of the multi-period solution* to the difference between i) the objective function value of the multi-period problem considering the static solution obtained as described above and ii) the optimal value of the multi-period problem.

We test both the weak and the strong value of the multi-period solution concepts just described on the CAB data set with 15 nodes for all the instances listed in Table 3. The random flow scenario (Scenario II) yields infeasibility with the static counterparts at some instances since it is not possible to build additional capacity modules in the static problems. Thus, we restrict our analysis on the value of the multi-period solution to the instances with increasing demand scenario (Scenario I). The results with both the single and multiple allocation problems are presented in Table 6.

For each instance, Table 6 lists i) the optimal solutions resulting from solving the proposed multi-period models with 5 periods, ii) the solutions obtained by solving the weak static counterpart, and iii) the solutions obtained by solving the strong static counterpart. For each instance, the locations of the hub nodes, the hub network state in each period, and total cost are presented. Total cost of the weak or strong value of the multi-period solution corresponds to total cost obtained using the data available for the whole planning horizon by fixing the corresponding solutions in the multi-period problem. The weak and strong relative values of the multi-period solution are also presented.

For all instances except one, the weak static solution turns out to be different from the optimal multi-period solution. Additionally, in 12 out of the 16 instances listed in Table 6, the weak static solution resulted in different hub locations when compared with the optimal multi-period solution. The highest percentage value of the weak value of the multi-period solution is 11.89% whereas the average is 4.30%. The results concerning the weak value of the multi-period solution show that the decision maker would endure additional costs during the planning horizon if the hub network is to be designed only by considering the worst case data and without considering the multi-period nature of the problem.

The percentage of the strong value of the multi-period solution are lower as expected. In 8 of the 16 instances listed in Table 6, the strong static counterpart led to the same solution as the optimal multi-period solution. All these instances correspond to the loose capacity set (Set II) where establishing a single capacity module in the beginning of the planning horizon can be sufficient to handle the flow in the remaining periods. In these instances, the multi-period solution resulted in a static network structure where neither the hub locations and capacities nor the resulting network change over time. In the remaining half of the instances, even though all the 5-period information is utilized, the strong static counterpart led to a different solution. Moreover, the locations of the hubs are different in 7 of the 8 instances in which the strong static counterpart led to a solution different from the multi-period one. The highest strong value of the multi-period solution (in percentage) is 5.63% with an average of 2.33% when the zeros are

Table 6: Value of the multi-period solution.

Instance number	5-period solution				Weak value of the multi-period solution				Strong value of the multi-period solution			
	Hub locations	Hub network state (C/I)	Total cost	Gap (%)	Hub locations	Hub network state (C/I)	Total cost	Gap (%)	Hub locations	Hub network state (C/I)	Total cost	Gap (%)
1	1,4,7,12	I-I-I-I-I	6884.63		1,4,6,7	I-I-I-I-I	7703.19	11.89	1,4,6,7,12	I-I-I-I-I	7041.61	2.28
2	1,4,7,12	I-I-I-I-I	7723.47		1,4,6,7	I-I-I-I-I	8267.55	7.04	1,4,6,7,12	I-I-I-I-I	7994.37	3.51
3	4,9,11	C-I-I-I-I	8287.19		4,6,7,13	I-I-I-I-I	8782.28	5.97	1,4,6,11	I-I-I-I-I	8702.49	5.01
4	4,9,11	C-I-I-I-I	8571.77		4,9,11,13	I-I-I-I-I	9168.33	6.96	4,5,9,11	I-I-I-I-I	9054.01	5.63
5	1,4,7,12	I-I-I-I-I	6838.73		4,7	C-C-C-C-C	7420.72	8.51	1,4,7,12	I-I-I-I-I	6838.73	0.00
6	4,11,12	I-I-I-I-I	7585.02		4,11	C-C-C-C-C	7845.82	3.44	4,11,12	I-I-I-I-I	7585.02	0.00
7	4,11	C-C-C-C-C	7962.85		4,11	C-C-C-C-C	8040.03	0.97	4,11	C-C-C-C-C	7962.85	0.00
8	4,11	C-C-C-C-C	8148.96		4,11	C-C-C-C-C	8234.24	1.05	4,11	C-C-C-C-C	8148.97	0.00
17	4,6,7,12	I-I-I-I-I	6739.00		5,11,12	I-I-I-I-I	7156.39	6.19	4,6,7,12	I-I-I-I-I	6750.75	0.17
18	4,7,12	I-I-I-I-I	7336.85		4,6,7	I-I-I-I-I	7532.75	2.67	4,6,7,12	I-I-I-I-I	7339.17	0.03
19	4,7	C-C-C-C-C	7696.17		4,6,7	I-I-I-I-I	7765.84	0.91	4,6,7	I-I-I-I-I	7736.36	0.52
20	4,7	C-C-C-C-C	7760.78		4,6,7	I-I-I-I-I	7874.50	1.47	4,6,7	I-I-I-I-I	7874.50	1.47
21	1,4,7,12	I-I-I-I-I	6744.55		4,7	C-C-C-C-C	7276.14	7.88	1,4,7,12	I-I-I-I-I	6744.55	0.00
22	4,7,12	I-I-I-I-I	7292.53		4,7	C-C-C-C-C	7475.33	2.51	4,7,12	I-I-I-I-I	7292.53	0.00
23	4,7	C-C-C-C-C	7660.70		4,7	C-C-C-C-C	7760.70	1.31	4,7	C-C-C-C-C	7660.70	0.00
24	4,7	C-C-C-C-C	7731.79		4,7	C-C-C-C-C	7731.79	0.00	4,7	C-C-C-C-C	7731.79	0.00

excluded, and an average of 1.16% including the zeros. Even though these percentages may seem low, the potential significance can be very high if total costs are high.

The analysis presented in Table 6 demonstrates the benefits of employing a multi-period modeling framework when designing hub networks, which gives strength to the multi-period models proposed in this paper.

5 Conclusion

In this paper, a modeling framework was proposed for multi-period hub location. We considered a pure phase-in problem in which a network is to be build from scratch. Nevertheless, we showed that our modeling framework can be easily adapted to the situation in which a network is already operating and the goal is to expand it over a finite planning horizon. We treated the single and multiple allocation cases separately. For each case, a mixed-integer linear programming formulation was proposed. Several sets of valid inequalities were also proposed for enhancing the models developed.

A set of computational tests performed with the models and their enhancements show that the hub networks are affected with the time dimension. The models can be tackled in reasonable time using a commercially available solver for moderately sized instances. For larger sized instances, there is a need for customized algorithms or heuristics which may be considered in future research.

Finally, we discussed the value of the multi-period solution which is a measure we propose for evaluating the relevance of using a multi-period modeling framework instead of a static counterpart. We proposed two different static counterparts for the multi-period problem and defined the weak and the strong value of the multi-period solution. The computational tests performed show that disregarding the multi-period nature of the problem may result in erroneously selected hub locations and the decision maker may have to endure additional costs.

This paper represents an attempt to obtain comprehensive modeling frameworks for multi-period hub location problems. Despite capturing several features of practical relevance, such framework can be extended or specialized in order to fit some requirements not investigated in this work. One such requirement concerns the possibility of establishing or expanding the capacity of a hub by installing more than one module in a location in each time period. Another feature which may be relevant when an existing network is being re-designed concerns the possibility of removing existing hubs or decreasing their operational capacity.

Finally, the results presented in this work show that from a methodological point of view much work can be developed. First, a deeper polyhedral analysis is required for finding additional

sets of valid inequalities that can enhance the models thus making the use of a general purpose solver more effective for obtaining optimal or near-optimal solutions to the problems. Second, large-scale instances of the problems investigated will hardly be tractable with such a solver. Accordingly, the development of tailored procedures (exact or approximate) is certainly another research direction worth exploring.

A Appendix

Table 7: Performance of the valid inequalities with single allocation problem.

Criteria	Instance number	Base model	(41)+(42)	(41)+(42) +(43)	Instance number	Base model	(41)+(42)	(41)+(42) +(43)
lp relaxation gap	1	5.94%	4.85%	4.36%	9	5.78%	4.57%	4.10%
root node gap		1.72%	1.47%	1.53%		1.40%	1.38%	1.34%
# of B&B nodes		12294	16588	10261		27	5493	3360
CPU time (sec)		1415.85	2025.00	934.90		650.40	651.07	503.35
lp relaxation gap	2	6.21%	5.68%	5.37%	10	6.06%	5.52%	5.22%
root node gap		2.50%	2.55%	2.49%		2.30%	2.42%	2.41%
# of B&B nodes		140770	83093	99925		50673	56172	93056
CPU time (sec)		19725.80	10608.30	18354.87		5012.50	4865.81	8972.04
lp relaxation gap	3	5.86%	5.64%	5.57%	11	5.62%	5.41%	5.35%
root node gap		3.50%	3.59%	3.52%		3.51%	3.45%	3.37%
# of B&B nodes		382663	246876	213316		209388	166334	217668
CPU time (sec)		42568.29	13410.75	13675.63		12203.49	11412.36	21130.09
lp relaxation gap	4	5.64%	5.53%	5.50%	12	5.51%	5.34%	5.32%
root node gap		3.82%	3.77%	5.04%		3.70%	3.80%	3.64%
# of B&B nodes		82979	127431	144816		45074	84118	87711
CPU time (sec)		5067.24	8015.18	7181.97		1661.93	4577.82	4433.36
lp relaxation gap	5	5.31%	4.22%	3.27%	13	5.39%	4.18%	3.24%
root node gap		0.60%	0.46%	0.56%		0.67%	0.81%	0.77%
# of B&B nodes		2603	2555	1543		26705	3972	20033
CPU time (sec)		435.20	421.02	308.79		1224.73	309.43	468.60
lp relaxation gap	6	4.56%	4.16%	3.55%	14	4.67%	4.25%	3.66%
root node gap		0.88%	0.86%	0.86%		1.04%	1.02%	1.02%
# of B&B nodes		2201	712	1330		7036	3657	2855
CPU time (sec)		284.86	422.73	385.06		992.68	499.34	320.24
lp relaxation gap	7	2.37%	2.23%	2.02%	15	2.51%	2.35%	2.16%
root node gap		0.40%	0.38%	0.42%		0.65%	0.66%	0.63%
# of B&B nodes		304	179	221		322	294	402
CPU time (sec)		120.59	86.11	105.11		126.07	104.40	86.67
lp relaxation gap	8	1.52%	1.26%	1.17%	16	1.75%	1.39%	1.30%
root node gap		0.10%	0.34%	0.37%		0.37%	0.59%	0.56%
# of B&B nodes		20	47	29		107	103	208
CPU time (sec)		31.61	30.02	23.54		37.99	31.12	45.55

Table 8: Performance of the valid inequalities with multiple allocation problem.

Criteria	Instance number	Base model	(41)+(42)	(41)+(42) +(43)	Instance number	Base model	(41)+(42)	(41)+(42) +(43)
lp relaxation gap	17	7.98%	6.51%	5.90%	25	7.75%	6.27%	5.69%
root node gap		1.78%	1.64%	1.63%		1.53%	1.43%	1.43%
# of B&B of nodes		1617	735	831		733	384	505
CPU time (sec)		1181.54	1101.62	1652.94		1227.87	893.56	1444.41
lp relaxation gap	18	9.53%	8.20%	7.75%	26	9.20%	7.84%	7.41%
root node gap		1.98%	1.85%	1.83%		1.72%	1.54%	1.57%
# of B&B nodes		489	306	936		211	117	1170
CPU time (sec)		1007.89	598.53	1271.81		809.29	527.49	1162.08
lp relaxation gap	19	10.87%	9.40%	9.11%	27	10.50%	9.01%	8.74%
root node gap		2.70%	2.17%	2.18%		2.38%	1.87%	1.87%
# of B&B nodes		2055	141	423		1915	553	187
CPU time (sec)		1038.84	450.20	698.10		933.81	509.39	684.88
lp relaxation gap	20	11.24%	9.81%	9.57%	28	10.84%	10.35%	10.11%
root node gap		2.60%	1.55%	1.57%		2.30%	1.31%	1.28%
# of B&B nodes		371	92	88		51	46	73
CPU time (sec)		251.22	195.22	296.68		217.09	222.29	415.10
lp relaxation gap	21	8.08%	6.64%	5.53%	29	7.77%	6.31%	5.25%
root node gap		1.49%	1.37%	1.36%		1.24%	1.09%	1.12%
# of B&B nodes		253	92	133		128	78	74
CPU time (sec)		1250.38	616.92	1063.13		892.34	594.05	954.09
lp relaxation gap	22	9.00%	7.67%	6.85%	30	8.66%	7.31%	6.50%
root node gap		1.11%	0.00%	0.96%		0.80%	0.00%	0.00%
# B&B nodes		25	0	17		25	0	0
CPU time (sec)		618.73	254.31	499.53		595.99	242.21	362.62
lp relaxation gap	23	10.47%	9.00%	8.43%	31	10.10%	8.61%	8.06%
root node gap		2.04%	1.55%	1.56%		1.71%	1.24%	1.25%
# of B&B nodes		823	525	39		47	15	59
CPU time (sec)		700.15	303.39	531.78		392.19	236.40	557.89
lp relaxation gap	24	10.91%	9.48%	8.99%	32	10.51%	9.05%	8.59%
root node gap		1.99%	1.10%	1.11%		1.70%	0.90%	0.91%
# of B&B nodes		33	11	11		36	9	7
CPU time (sec)		203.91	157.34	245.01		217.78	118.25	193.29

References

- [1] ALUMUR, S., AND KARA, B. Network hub location problems: The state of the art. *European Journal of Operational Research* 190 (2008), 1–21.
- [2] ALUMUR, S., KARA, B., AND KARASAN, O. The design of single allocation incomplete hub networks. *Transportation Research Part B* 43 (2009), 936–951.
- [3] ALUMUR, S., KARA, B., AND KARASAN, O. Multimodal hub location and hub network design. *Omega* 40 (2012), 927–939.
- [4] ALUMUR, S., NICKEL, S., SALDANHA-DA-GAMA, F., AND VERTER, V. Multi-period reverse logistics network design. *European Journal of Operational Research* 220 (2012), 67–78.
- [5] BEASLEY, J. E. OR Library: Hub location. <http://people.brunel.ac.uk/mastjjb/jeb/orlib/phubinfo.html>, 1990.
- [6] BOLAND, N., KRISHNAMOORTHY, M., ERNST, A., AND EBERY, J. Preprocessing and cutting for multiple allocation hub location problems. *European Journal of Operational Research* 155 (2004), 638–653.
- [7] CAMPBELL, J. Locating Transportation Terminals To Serve An Expanding Demand. *Transportation Research Part B* 24 (1990), 173–192.
- [8] CAMPBELL, J., ERNST, A., AND KRISHNAMOORTHY, M. *Facility Location, Applications and Theory*. Springer, 2002, ch. Hub Location Problems.
- [9] CAMPBELL, J. F., AND O’KELLY, M. E. Twenty-five years of hub location research. *Transportation Science* 46 (2012), 153–169.
- [10] CONTRERAS, I., CORDEAU, J.-F., AND LAPORTE, G. The dynamic uncapacitated hub location problem. *Transportation Science* 45 (2011), 18–32.
- [11] CONTRERAS, I., FERNÁNDEZ, E., AND MARÍN, A. The tree of hubs location problem. *European Journal of Operational Research* 202 (2010), 390–400.
- [12] CORDEAU, J.-F., PASIN, F., AND SOLOMON, M. M. An integrated model for logistics network design. *Annals of Operations Research* 144 (2006), 59–82.

- [13] CORREIA, I., NICKEL, S., AND SALDANHA-DA-GAMA, F. Single-assignment hub location problems with multiple capacity levels. *Transportation Research Part B* 44 (2010), 1047–1066.
- [14] EBERY, J., KRISHNAMOORTHY, M., ERNST, A., AND BOLAND, N. The capacitated multiple allocation hub location problem: Formulations and algorithms. *European journal of Operational Research* 120 (2000), 614–631.
- [15] ERNST, A., AND KRISHNAMOORTHY, M. Solution algorithms for the capacitated single allocation hub location problem. *Annals of Operations Research* 86 (1999), 141–159.
- [16] ERNST, A. T., AND KRISHNAMOORTHY, M. Efficient algorithms for the uncapacitated single allocation p-hub median problem. *Location Science* 4 (1996), 139–154.
- [17] ERNST, A. T., AND KRISHNAMOORTHY, M. Exact and heuristic algorithms for the uncapacitated multiple allocation p-hub median problem. *European Journal of Operational Research* 104 (1998), 100–112.
- [18] GELAREH, S. *Hub Location Models in Public Transport Planning*. PhD thesis, Vom Fachbereich Mathematik der Technischen Universität Kaiserslautern, 2008.
- [19] GELAREH, S., AND NICKEL, S. Hub location problems in transportation networks. *Transportation Research Part E* 47 (2011), 1092–1111.
- [20] MELO, M., NICKEL, S., AND SALDANHA-DA-GAMA, F. Dynamic multi-commodity capacitated facility location: a mathematical modeling framework for strategic supply chain planning. *Computers & Operations Research* 33 (2006), 181–208.
- [21] MELO, M., NICKEL, S., AND SALDANHA-DA-GAMA, F. Facility location and supply chain management a review. *European Journal of Operational Research* 196 (2009), 1–12.
- [22] NICKEL, S., SCHÖBEL, A., AND SONNEBORN, T. *Mathematics Methods and Optimization in Transportation Systems*. Springer, 2001, ch. Hub Location Problems in Urban Traffic Networks.
- [23] O’KELLY, M. A quadratic integer problem for the location of interacting hub facilities. *European Journal of Operational Research* 32 (1987), 393–404.
- [24] TAN, P. Z., AND KARA, B. Y. A hub covering model for cargo delivery systems. *Networks* 49 (2007), 28–39.

- [25] YAMAN, H. Star p-hub median problem with modular arc capacities. *Computers & Operations Research* 35, 9 (2008), 3009–3019.
- [26] YOON, M.-G., AND CURRENT, J. The hub location and network design problem with fixed and variable costs: formulation and dual-based solution heuristic. *Journal of the Operational Research Society* 59 (2008), 80–89.