Centro de Investigação Operacional

Multi-period hub location problems in transportation networks

Isabel Correia, Shahin Gelareh, Stefan Nickel and
Francisco Saldanha-da-Gama

CIO – Working Paper 9/2012

Multi-period hub location problems in transportation networks

Isabel Correia[a], Shahin Gelareh[b*], Stefan Nickel[c,d],
Francisco Saldanha-da-Gama[e]

[a] *Departamento de Matemática - CMA, Faculdade de Ciências e Tecnologia, Universidade Nova Lisboa, 2829-516 Caparica, Portugal*
[b] *Department of Management Engineering, Technical University of Denmark, Produktionstorvet, Building 426, 2800 Kgs. Lyngby, Denmark*
[c] *Institute for Operations Research, University of Karlsruhe (TH), Karlsruhe, Germany*
[d] *Fraunhofer Institute for Industrial Mathematics (ITWM), Kaiserslautern, Germany*
[e] *DEIO-CIO, Faculdade de Ciências, Universidade de Lisboa, 1716-049 Lisboa, Portugal*

December 21, 2012

Abstract

In this paper we consider a phase-in/phase-out hub location problem. In particular, we assume that a transportation system is currently operating which can be changed in the future. A planning horizon partitioned into several consecutive time period is considered. Changes in the network structure can occur in the hubs and in the hub edges. More specifically, existing elements can be removed and new elements can be established. The problem consists in deciding, which network structure should be operating in each period of the planning horizon and how the flow should be routed through the operating structure in order to minimize the total cost. In each period, a budget is considered for making changes in the network structure. For this problem, a mixed-integer linear programming problem is proposed. Due to the complexity of the problem, a local search based heuristic is also developed. The results of a series of extensive computational tests are reported.

Keywords: Hub Location, Transportation Networks, Multi-Period Network Design, Heuristics.

*Corresponding author. *E-mail address*: shahin.gelareh@gmail.com

1 Introduction

The design of hub networks is a problem of great relevance in the fields of transportation and telecommunications. The typical structure of these networks includes a set of nodes - called hubs - that consolidate the flow, and a set of edges - called hub edges -, which connect the hubs. The non-hub nodes are assigned to one or more hubs so that all the flow originated at the nodes is routed via at least one hub. Often, no direct shipments are allowed between non-hub nodes. This type of structure can be very attractive because it often allows important savings in the routing costs due to economies of scale that result from flow consolidation at the hubs as well as from flow routing through hub edges. For instance, the use of hub edges leads to gains in the performance of the system by the possibility of making use of larger vehicles.

In a hub location problem (HLP) the goal is to select the nodes that should become hubs and also to decide how to allocate the non-hub nodes to the hubs so that the total cost for building and operating the network is minimized (see, for instance, Campbell et al. [11]). Two possibilities are often considered for allocating the non-hub nodes to the hubs: single or multiple allocation. In the former, each non-hub node is allocated to exactly one hub; in the latter, this is not the case. Depending on the existence of capacities, a HLP can be classified as capacitated or uncapacitated.

Since the seminal paper by O'Kelly [35] much research has been devoted to hub location problems. In particular, many different hub location problems have been addressed in the literature. O'Kelly [34] proposed the first mathematical programming formulation for a hub location problem in a discrete setting, namely for the single-allocation p-hub median problem. In such problem, the number of hubs to be installed is specified in advance. This problem and variants of it have been studied by many authors (e.g. Campbell [9], Skorin-Kapov et al. [37], Ernst and Krishnamoorthy [22, 23]).

An alternative to p-hub location problems is to leave free the number of hubs to install. In most of the situations, a fixed setup cost for the hubs is considered. Capacitated and uncapacitated versions of these variants have been addressed by many authors such as Campbell [9], Abdinnour-Helm [1], Abdinnour-Helm and Venkataramanan [2], Correia et al. [20], Ernst and Krishnamoorthy [24], Ebery et al. [21], Mayer and Wagner [31], Boland et al. [7], Labbé et al. [30], Topcuoglu et al. [38], Chen [14], and Contreras et al. [16].

Other hub location problems include the one studied by Yaman and Carello [40] in which modular link capacities are considered and the problem studied by Correia et al. [19], who introduce capacity decisions in a single-allocation hub location problem.

In many hub location problems addressed by the existing literature, the following assumptions are considered:

1. The hub level network is a complete graph.
2. There is a discount factor associated with the use of hub edges, which reflects economies of scale.
3. Direct connections between non-hub nodes are not allowed.
4. Costs are proportional to the distances i.e, the triangle inequality holds.
5. All nodes are candidates to become hubs.

Note that assumptions 1, 3 and 4 together imply that all flow is routed via exactly 1 or 2 hubs. The reader can refer to Nickel et al. [33] for a deeper discussion on these aspects.

As mentioned by Alumur and Kara [4], the previous assumptions hold in many practical applications such as in computer networks, postal-delivery, less-than-truck loading and supply chain management. Nevertheless, in other cases, some of the above conditions are too restrictive. For instance, 1) has been relaxed in several contexts such as urban public transport networks (Nickel et al. [33]), freight transportation (Alumur and Kara [5, 6]) and telecommunications (e.g. Contreras et al. [17, 18]). Campbell et al. [12, 13] and Yaman [39] also address problems with incomplete hub networks.

In practice, one may have to relax other of the above conditions. One example regards the potential nodes to become hubs. In real-life problems it often happens that not all locations have the necessary conditions (e.g. space requirements or accessibility) to become a hub.

One feature that cannot be ignored in many problems regards the existence of a system already operating - even if rudimentary. Examples can be found in freight transportation, liner shipping and city railway networks. Accordingly, the plan to develop should take this fact into account by considering the possibility of removing some existing hubs or hub edges and establishing new ones.

The establishment of new structures as well as the removal of existing ones are often medium or long-term projects involving time-consuming activities (e.g. construction) and capital investment (e.g. installation of an adequate infrastructure, equipment supply, and employee training). For instance, building a new underground line is quite time consuming and is often done progressively in different stages. Building a transhipment point for a cargo company is another example. In addition to this, the task of planning a network has often to be made in such a way that the implementation of changes in the network structure are carried out smoothly without disrupting the network flows. Moreover, to abate the financial burden put on the company by such projects it is often the case that capital expenditures as well as network design decisions should be planned over several time periods. For instance, public transportation systems are highly dependent of state or municipal funding, which is often not available at once but in different moments in time. In the case of a private company, large investment are often split over a set of periods of time (e.g. years). Last but not least, the flows between origin/destination pairs and the costs often change over time. This is the case in many (not to say all) transportation systems.

The previous aspects suggest the use of a dynamic plan in opposition to a static one. One possibility for achieving this consists in considering a planning horizon divided into several time periods and then accept that changes in the network structure can be made at the transition between consecutive time periods.

In this paper we address a multi-period hub location problem (MPHLP) considering that a network may already be operating. Changes in the network structure are allowed during the planning horizon. Such changes refer to the removal of existing hubs and hub edges and the establishment of new ones. Once a hub or a hub edge is removed, it can not be established again and conversely, once a new hub or hub edge is installed, it should remain operating until the end of the planning horizon. We assume that the hub level network is connected but not necessarily a clique. An exogenous budget is considered in each period for installing and removing hubs and hub edges as well as for maintaining the operating ones. The budget available but not used in some period can be invested and its return used in subsequent periods. The flows between the origin destination pairs may vary over the planning horizon and are consolidated at the hubs under a multiple allocation pattern.

The problem consists in choosing the hub level network structure that should be operating in each period so that the total cost over the planning horizon is minimized. By hub level network structure in each period we mean the hubs and hub edges that are operating. Several cost components are considered: a) the cost for establishing new hubs and hub-edges (e.g. equipment construction and land acquisition); b) the cost for removing existing hubs and hub-edges (e.g. workforce transfer); c) the cost for operating hubs and hub-edges; d) the cost for routing the flows. A discount factor is considered for the flow routed through the hub edges which reflect the economies of scale mentioned above.

Figure 1 illustrates a 6-node network with a design that changes over 3 time periods. In this figure, the square nodes identify hubs. Multiple allocation is allowed (e.g. node k in period 1). In the situation depicted in this Figure, node k becomes a hub in period 2 whereas node m, which was a hub in the beginning, becomes a non-hub node in period 3.

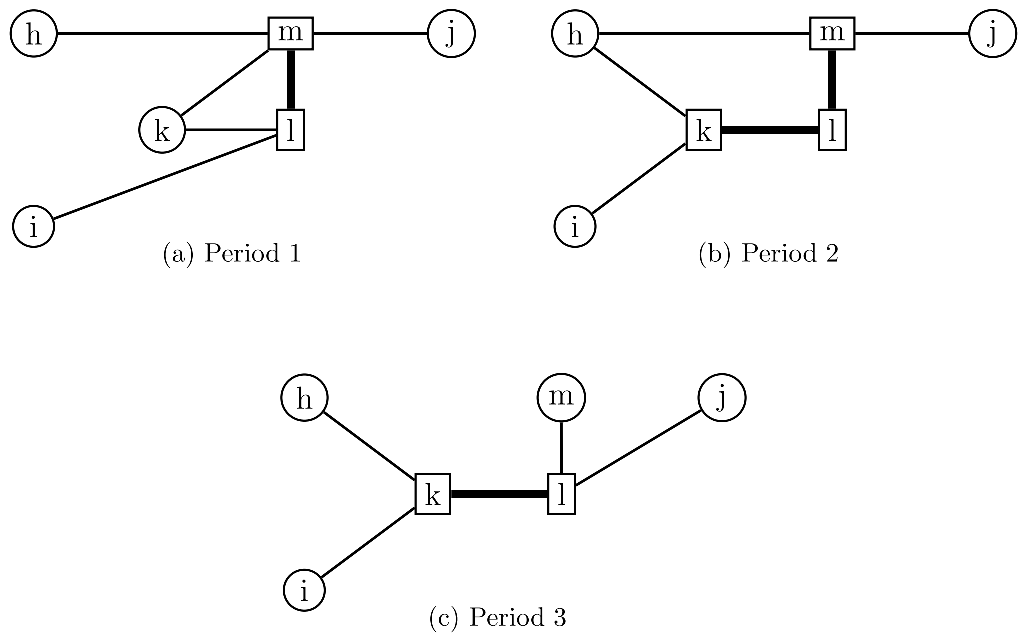


Figure 1: Example of a network evolving with the time.

For the problem just described and illustrated we propose a mixed-integer linear programming formulation. The complexity of the problem prevents it from being solved to optimality using a general solver unless toy instances are considered. For this reason, we propose a heuristic for obtaining feasible solutions to the problem.

The first paper addressing a hub location problem in a dynamic context is due to

Campbell [8]. A continuous approximation model of a general freight carrier serving a fixed region with an increasing density of demand is considered. Terminal, transportation and relocation costs are considered. As the demand for service by a freight carrier increases, transportation terminals are added in order to decrease the transportation costs. A myopic strategy with limited capability for relocation is shown to be nearly optimal unless terminal relocation costs are large.

Gelareh [25] propose the first models for multi-period hub location problems in the context of public transportation systems. Several variants are considered, including ones with a constraint on the number of facilities and in the available budget for changing the network design. Mixed-integer linear programming formulations are proposed for both situations. Nevertheless, only for very small instances can the problem be solved to optimality. For this reason, a heuristic procedure is also proposed in order to obtain feasible solutions to the problem.

More recently, Contreras et al. [15] consider a multi-period uncapacitated hub location problem in which the location of the hubs can change over time in order to cope with a set of flows changing over time between the origin/destination pairs. A planning horizon is considered which is divided into several time periods. In each period, all the demand should be routed through the network. The total cost over the planning horizon is to be minimized and includes the costs for locating, operating and closing hubs as well as the cost of routing the flows. A quadratic integer programming formulation is proposed. A Lagrangean relaxation approach is proposed which is embedded in a tree search procedure in order to obtain the optimal solution to the problem.

Despite the small number of papers addressing dynamic hub location problems, the importance of considering models capturing the dynamic nature of some facility location problems is clear in the literature. The reader can refer, for instance, to the works by Albareda-Sambola et al. [3], Melo et al. [32] and the references therein.

As we mentioned above, in this paper, we also propose a heuristic approach. Due to the complexity of many hub location problems, several authors have proposed heuristic procedures for obtaining feasible solutions.

Simulated Annealing procedures were proposed by (Ernst and Krishnamoorthy [22, 24]), whereas Tabu Search approaches were proposed by (Klincewicz [29], Skorin-Kapov et al. [36]). Abdinnour-Helm and Venkataramanan [2], and Topcuoglu et al. [38] considered the use of Genetic Algorithms. GRASP was attempted by Klincewicz [29]. We can also find hybrid approaches in the literature which is the case with the works by Abdinnour-Helm [1] and Chen [14]. Some other heuristic approaches that cannot be classified in any of the above categories regard the works by Klincewicz [28], Campbell [10], Ernst and Krishnamoorthy [23], Ebery et al. [21] and Gelareh and Nickel [26]. For a more comprehensive review on the problems and applications that have been addressed in the literature the reader can refer to the review by Alumur and Kara [4] and to the book chapters by Campbell et al. [11] and Kara and Taner [27].

The remainder of this paper is organized as follows. In section 2, a mathematical programming formulation is proposed for the studied problem. In section 3 a heuristic approach is proposed. The results of the computational experiments conducted are presented and analyzed in section 4. The paper ends with the conclusions drawn from the work done.

2 Mathematical programming formulation

In this section we start by introducing the notation that will be used throughout the paper and afterwards we propose and discuss a mixed-integer linear programming formulation.

As we mentioned in section 1, we assume that a network may already be in operation and thus the problem is to find the best way to change it over time. We call the *initial configuration* the set of hubs and hub edges that are in operation before the beginning of the planning horizon. These are the components of the network that can be removed during the planning horizon. Recall that according to the definition presented in the previous section, the end points of a hub edge are operating hub nodes. Moreover, recall our assumption stating that the network defined by the hubs and hub edges is always connected.

2.1 Notation

Sets

$\mathcal{T} = \{1, ..., T\}$	Set of time periods in the planning horizon.
$N = \{1, ..., n\}$	Set of nodes.
H^c	Set of hubs in the initial configuration which can be removed during the planning horizon.

H^o	Set of non-hub nodes in the initial configuration that can become hubs during the planning horizon.
$H = H^o \cup H^c$	
E^c	Set of hub-edges $\{k, l\}$, $k < l$, in the initial configuration which can be removed during the planning horizon.
E^o	Set of non-hub edges $\{k, l\}$, $k < l$, in the initial configuration that can become hub-edges during the planning horizon.
$E = E^o \cup E^c$	

Costs

FO_k^t	Set-up cost incurred for establishing node k as a hub in the beginning of period t, $k \in H^o$, $t \in \mathcal{T}$.
FC_k^t	Cost for closing the existing hub node k at the end of period t, $k \in H^c$, $t \in \mathcal{T} \setminus T$.
FM_k^t	Maintenance cost for hub k in period t, $k \in H$, $t \in \mathcal{T}$.
GO_{kl}^t	Set-up cost incurred for establishing the hub edge $\{k, l\}$ in the beginning of period t, $\{k, l\} \in E^o$, $t \in \mathcal{T}$.
GC_{kl}^t	Cost for closing hub edge $\{k, l\}$ at the end of period t, $\{k, l\} \in E^c$, $t \in \mathcal{T} \setminus T$.
GM_{kl}^t	Maintenance cost for the hub edge $\{k, l\}$ in period t, $\{k, l\} \in E$, $t \in \mathcal{T}$.

We define:

$$F_k^t = FO_k^t + \sum_{\tau=t}^{T} FM_k^\tau, \quad k \in H^o, t \in \mathcal{T}$$

$$F_k^t = FC_k^t + \sum_{\tau=1}^{t} FM_k^\tau, \quad k \in H^c, t \in \mathcal{T} \setminus T$$

$$F_k^T = \sum_{\tau=1}^{T} FM_k^\tau, \quad k \in H^c$$

$$G_{kl}^t = GO_{kl}^t + \sum_{\tau=t}^{T} GM_{kl}^\tau, \quad \{k, l\} \in E^o, t \in \mathcal{T}$$

$$G_{kl}^t = GC_{kl}^t + \sum_{\tau=1}^{t} GM_{kl}^\tau, \quad \{k, l\} \in E^c, t \in \mathcal{T} \setminus T$$

$$G_{kl}^T = \sum_{\tau=1}^{T} GM_{kl}^{\tau}, \quad \{k,l\} \in E^c$$

The costs F and G simply give the total cost that is incurred by a hub or a hub edge (maintenance for all periods in which the hub or hub edge is operating plus set-up or removal depending on the situation).

Other parameters

W_{ij}^t	Flow to be sent from node i to node j in period t, $i,j \in N$, $t \in \mathcal{T}$.
C_{ij}^t	Cost for sending one unit of flow through edge $\{i,j\}$ in period t, $i,j \in N$, $t \in \mathcal{T}$.
α^t	Discount factor for using a connection between hubs in period t, $t \in \mathcal{T}$. It is assumed that $0 < \alpha^t < 1$, $t \in \mathcal{T}$.
B^t	Exogenous budget available at the beginning of time period t, $t \in \mathcal{T}$.
ρ^t	Unit return rate for the capital not invested in period t, $t \in \mathcal{T}$.
η^t	Available budget at the end of period t, $t \in \mathcal{T}$.

It is assumed that the cost matrix $[C_{ij}^t]_{i,j \in N}$ is symmetric ($t \in \mathcal{T}$) and also that $C_{ii}^t = 0$ ($i \in N$, $t \in \mathcal{T}$). Regarding the flow matrix $[W_{ij}^t]_{i,j \in N}$ for each $t \in \mathcal{T}$, we assume that $W_{ii}^t = 0$, $i \in N$.

Regarding the budget available, we assume that it allows the initial configuration to be maintained over the entire planning horizon i.e, we assume that the budget is enough to allow the implementation of such solution.

2.2 Mixed-Integer linear programming formulation

We consider two types of decision variables: network design decision variables, which define the hubs and hub edges operating in each period, and flow decision variables that define how the flow is routed through the network in each period.

Network design decision variables:

For $k \in H^o$, $t \in \mathcal{T}$:

$$y_k^t = \begin{cases} 1 & \text{if node } k \text{ is established as a hub node at the beginning of period } t \\ 0 & \text{otherwise} \end{cases}$$

For $k \in H^c$, $t \in \mathcal{T} \setminus T$:

$$y_k^t = \begin{cases} 1 & \text{if existing hub node } k \text{ finishes its operation as a hub at the end of period } t \\ 0 & \text{otherwise} \end{cases}$$

For $k \in H^c$:

$$y_k^T = \begin{cases} 1 & \text{if existing hub node } k \text{ operates over the entire planning horizon} \\ 0 & \text{otherwise} \end{cases}$$

For $\{k, l\} \in E^o$, $t \in \mathcal{T}$:

$$z_{kl}^t = \begin{cases} 1 & \text{if edge } \{k, l\} \text{ is established as a hub edge at the beginning of period } t \\ 0 & \text{otherwise} \end{cases}$$

For $\{k, l\} \in E^c$, $t \in \mathcal{T} \setminus T$:

$$z_{kl}^t = \begin{cases} 1 & \text{if the existing hub edge } \{k, l\} \text{ finishes its operation as a hub edge at the} \\ & \text{end of period } t \\ 0 & \text{otherwise} \end{cases}$$

For $\{k, l\} \in E^c$:

$$z_{kl}^T = \begin{cases} 1 & \text{if the existing hub edge } \{k, l\} \text{ operates during the entire planning horizon} \\ 0 & \text{otherwise} \end{cases}$$

Flow decision variables:

x_{ijkl}^t : Fraction of the flow from i to j in period t that is routed via the hub edge $\{k, l\}$ in the direction $k \to l$, $i, j, k, l \in N$, $i \neq j$, $k \neq l$, $t \in \mathcal{T}$.

u_{ijk}^t : Fraction of the flow from non-hub origin i to destination j in period t that leaves i via hub k, $i, j, k \in N$, $i \neq j$, $k \neq i, j$, $t \in \mathcal{T}$.

v^t_{ijl} : Fraction of the flow sent from node i to non-hub node j that arrives at j via hub l, $i, j, l \in N$, $i \neq j$, $l \neq i, j$, $t \in \mathcal{T}$.

s^t_{ij} : Fraction of the flow from node i to node j in period t that traverses the non-hub edge $\{i, j\}$ in the direction $i \to j$, $i, j \in N$, $i \neq j$, $t \in \mathcal{T}$.

Figure 2 illustrates the flow decision variables defined above. Note that the u-variables and v-variables are similar to the flow variables proposed by Ernst and Krishnamoorthy [23].

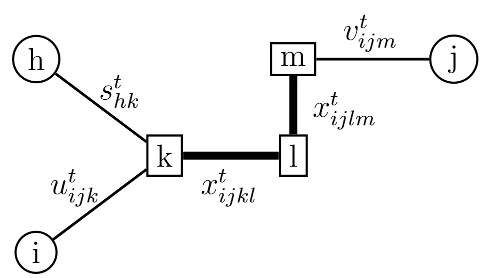


Figure 2: Flow variables in period $t \in \mathcal{T}$.

For the sake of readability we present a MILP formulation for MPHLP using several categories of constraints as follows:

1. Flow collection, distribution and conservation.
2. The initial configuration should hold for the first period.
3. Consistency between operating hubs and operating hub-edges.
4. Consistency between the flow traversing the hub edges and the operating hub edges.
5. Consistency between flow traversing non-hub edges and the actual non-hub edges.
6. Consistency between flow starting in the non-hub edges and the actual non-hub edges.
7. Consistency between flow ending in the non-hub edges and the actual non-hub edges.
8. Consistency between the flow traversing the hub nodes and the actual operating hub nodes.

9. Budget constraints.

10. Variable fixing.

1. Flow collection, distribution and conservation

- In every period of the planning horizon, the flow from i to j (j to i) leaves (arrives at) node i using a hub edge or a non-hub edge:

$$\sum_{l \neq i} x^t_{ijil} + \sum_{l \neq i,j} u^t_{ijl} + s^t_{ij} = 1 \qquad i, j \in N,\ i < j,\ t \in \mathcal{T} \tag{1}$$

- In every time period, the flow from i to j (j to i) arrives at (leaves) j via a hub edge or a non-hub edge:

$$\sum_{l \neq j} x^t_{ijlj} + \sum_{l \neq i,j} v^t_{ijl} + s^t_{ij} = 1 \qquad i, j \in N,\ i < j,\ t \in \mathcal{T} \tag{2}$$

- In every period, the flow from i to j (j to i) is not lost in the hub level network:

$$\sum_{l \neq i,k} x^t_{ijkl} + v^t_{ijk} = \sum_{l \neq j,k} x^t_{ijlk} + u^t_{ijk} \quad i, j, k \in N,\ i < j,\ k \neq i, j,\ t \in \mathcal{T} \tag{3}$$

2. The initial configuration should hold for the first period

- The existing hubs must operate at least in period 1:

$$\sum_{t \in \mathcal{T}} y^t_k = 1 \qquad k \in H^c \tag{4}$$

- The existing hub-edges must operate at least in period 1:

$$\sum_{t \in \mathcal{T}} z^t_{kl} = 1 \qquad \{k, l\} \in E^c \tag{5}$$

3. Consistency between operating hubs and operating hub edges

- In order to establish a new hub edge in some time period the corresponding extreme nodes should be hubs operating at least in the same time periods implied by the establishment of the hub edge. Accordingly,

$$z_{kl}^t \leq \sum_{\tau=1}^{t} y_k^\tau, \qquad \{k,l\} \in E^o,\ k \in H^o,\ l \in H,\ t \in \mathcal{T} \tag{6}$$

$$z_{kl}^t \leq y_k^T, \qquad \{k,l\} \in E^o,\ k \in H^c,\ l \in H,\ t \in \mathcal{T} \tag{7}$$

$$z_{kl}^t \leq \sum_{\tau=1}^{t} y_l^\tau, \qquad \{k,l\} \in E^o,\ k \in H,\ l \in H^o,\ t \in \mathcal{T} \tag{8}$$

$$z_{kl}^t \leq y_l^T, \qquad \{k,l\} \in E^o,\ k \in H,\ l \in H^c,\ t \in \mathcal{T} \tag{9}$$

- In each time period, an existing hub edge can only be operating if both end nodes (which correspond to hubs operating in the initial configuration) have not been removed so far.

$$z_{kl}^t \geq y_k^t, \qquad \{k,l\} \in E^c,\ k \in H^c,\ t \in \mathcal{T} \setminus T \tag{10}$$

$$z_{kl}^t \geq y_l^t, \qquad \{k,l\} \in E^c,\ l \in H^c,\ t \in \mathcal{T} \setminus T \tag{11}$$

4. Consistency between the flow traversing the hub edges and the operating hub edges

- Flow can only traverse a potential hub edge in some time period if the edge has been installed.

$$x_{ijkl}^t + x_{ijlk}^t \leq \sum_{\tau=1}^{t} z_{kl}^\tau, \qquad \{k,l\} \in E^o,\ i,j \in N,\ i < j,\ t \in \mathcal{T} \tag{12}$$

- Flow can only traverse an existing hub edge in some time period if the edge has not been removed.

$$x_{ijkl}^t + x_{ijlk}^t \leq 1 - \sum_{\tau=1}^{t-1} z_{kl}^\tau, \qquad \{k,l\} \in E^c,\ i,j \in N,\ i < j,\ t \in \mathcal{T} \tag{13}$$

5. Consistency between flow traversing non-hub edges and the actual non-hub edges

A variable s_{ij}^t can only be greater than 0 if edge $\{i,j\}$ is a non-hub edge in period t. Accordingly, some consistency constraints are needed. Four cases must be distinguished:

$$s_{ij}^t \leq \left| \sum_{\tau=1}^{t} \left(y_i^\tau - y_j^\tau \right) \right|, \qquad i,j \in H^o,\ i<j,\ t \in \mathcal{T}$$

$$s_{ij}^t \leq 1 - \left| \sum_{\tau=1}^{t} y_i^\tau - \sum_{\tau=1}^{t-1} y_j^\tau \right|, \qquad i \in H^o,\ j \in H^c,\ i<j,\ t \in \mathcal{T}$$

$$s_{ij}^t \leq 1 - \left| \sum_{\tau=1}^{t-1} y_i^\tau - \sum_{\tau=1}^{t} y_j^\tau \right|, \qquad i \in H^c,\ j \in H^o,\ i<j,\ t \in \mathcal{T}$$

$$s_{ij}^t \leq \left| \sum_{\tau=1}^{t-1} \left(y_i^\tau - y_j^\tau \right) \right|, \qquad i,j \in H^c,\ i<j,\ t \in \mathcal{T}$$

The previous constraints can be easily linearized as follows by considering additional variables $\delta_{ij}^+ \geq 0$ and $\delta_{ij}^- \geq 0$.

$$\sum_{\tau=1}^{t} \left(y_i^\tau - y_j^\tau \right) = \delta_{ij}^{t+} - \delta_{ij}^{t-}, \qquad i,j \in H^o,\ i<j,\ t \in \mathcal{T} \tag{14}$$

$$1 - \sum_{\tau=1}^{t} y_i^\tau - \sum_{\tau=1}^{t-1} y_j^\tau = \delta_{ij}^{t+} - \delta_{ij}^{t-}, \qquad i \in H^o,\ j \in H^c,\ i<j,\ t \in \mathcal{T} \tag{15}$$

$$1 - \sum_{\tau=1}^{t-1} y_i^\tau - \sum_{\tau=1}^{t} y_j^\tau = \delta_{ij}^{t+} - \delta_{ij}^{t-}, \qquad i \in H^c,\ j \in H^o,\ i<j,\ t \in \mathcal{T} \tag{16}$$

$$\sum_{\tau=1}^{t-1} \left(y_i^\tau - y_j^\tau \right) = \delta_{ij}^{t+} - \delta_{ij}^{t-}, \qquad i,j \in H^c,\ i<j,\ t \in \mathcal{T} \tag{17}$$

$$s_{ij}^t \leq \delta_{ij}^{t+} + \delta_{ij}^{t-} \leq 1, \qquad i,j \in H,\ i<j,\ t \in \mathcal{T} \tag{18}$$

6. Consistency between flow starting in the non-hub edges and the actual non-hub edges

A variable u_{ijk}^t can only be greater than 0 if node i is not an operating hub.

$$u_{ijk}^t \leq 1 - \sum_{\tau=1}^{t} y_i^\tau, \qquad i \in H^o,\ i<j,\ k \neq i,j,\ t \in \mathcal{T} \tag{19}$$

$$u_{ijk}^t \leq \sum_{\tau=1}^{t-1} y_i^\tau, \qquad i \in H^c,\ i<j,\ k \neq i,j,\ t \in \mathcal{T} \tag{20}$$

7. Consistency between flow ending in the non-hub edges and the actual non-hub edges

A variable v_{ijl}^t can only be greater than 0 if node j is not an operating hub.

$$v_{ijl}^t \leq 1 - \sum_{\tau=1}^{t} y_j^\tau, \qquad j \in H^o,\ i<j,\ k \neq i,j,\ t \in \mathcal{T} \tag{21}$$

$$v_{ijl}^t \leq \sum_{\tau=1}^{t-1} y_j^\tau, \qquad j \in H^c,\ i<j,\ k \neq i,j,\ t \in \mathcal{T} \tag{22}$$

8. Consistency between the flow traversing the hubs and the actual operating hubs

Flow can only traverse a hub node if the node corresponds to an operating hub.

$$u_{ijk}^t + \sum_{l \neq j,k} x_{ijlk}^t \leq \sum_{\tau=1}^{t} y_k^\tau, \qquad i,j \in N,\ i<j,\ k \in H^o,\ k \neq i,j,\ t \in \mathcal{T} \tag{23}$$

$$u_{ijk}^t + \sum_{l \neq j,k} x_{ijlk}^t \leq 1 - \sum_{\tau=1}^{t-1} y_k^\tau, \qquad i,j \in N,\ i<j,\ k \in H^c,\ k \neq i,j,\ t \in \mathcal{T} \tag{24}$$

$$v_{ijl}^t + \sum_{k \neq i,l} x_{ijlk}^t \leq \sum_{\tau=1}^{t} y_l^\tau, \qquad i,j \in N,\ i<j,\ l \in H^o,\ l \neq i,j,\ t \in \mathcal{T} \tag{25}$$

$$v_{ijl}^t + \sum_{k \neq i,l} x_{ijlk}^t \leq 1 - \sum_{\tau=1}^{t-1} y_l^\tau, \qquad i,j \in N,\ i<j,\ l \in H^c,\ l \neq i,j,\ t \in \mathcal{T} \tag{26}$$

$$s_{ij}^t + 2x_{ijij}^t + \sum_{l \neq i,j} \left(x_{ijil}^t + x_{ijlj}^t\right) \leq \begin{cases} \sum_{\tau=1}^{t} \left(y_i^\tau + y_j^\tau\right), & i,j \in H^o,\ i<j,\ t \in \mathcal{T} \\ \sum_{\tau=1}^{t} y_i^\tau + \left(1 - \sum_{\tau=1}^{t-1} y_j^\tau\right), & i \in H^o,\ j \in H^c,\ i<j,\ t \in \mathcal{T} \\ \left(1 - \sum_{\tau=1}^{t-1} y_i^\tau\right) + \sum_{\tau=1}^{t} y_j^\tau, & i \in H^c, : j \in H^o,\ i<j,\ t \in \mathcal{T} \\ 2 - \left(\sum_{\tau=1}^{t-1} y_i^\tau + \sum_{\tau=1}^{t-1} y_j^\tau\right), & i,j \in H^c,\ i<j,\ t \in \mathcal{T} \end{cases} \tag{27}$$

9. Budget constraints

In each time period, the amount spent with the hub level network (establishments, removals and maintenances) cannot exceed the existing budget.

$$\sum_{k\in H^o}\left(FO_k^1+FM_k^1\right)y_k^1+\sum_{k\in H^c}\left(FC_k^1y_k^1+FM_k^1\right)+\\ \sum_{\{k,l\}\in E^o}\left(GO_{kl}^1+GM_{kl}^1\right)z_{kl}^1+\sum_{\{k,l\}\in E^c}\left(GC_{kl}^1z_{kl}^1+GM_{kl}^1\right)+\eta^1=B^1 \quad (28)$$

$$\sum_{k\in H^o}\left(FO_k^ty_k^t+FM_k^t\sum_{\tau=1}^{t}y_k^\tau\right)+\sum_{k\in H^c}\left(FC_k^ty_k^t+FM_k^t\left(1-\sum_{\tau=1}^{t-1}y_k^\tau\right)\right)+\\ \sum_{\{k,l\}\in E^o}\left(GO_{kl}^tz_{kl}^t+GM_{kl}^t\sum_{\tau=1}^{t}z_{kl}^\tau\right)+\sum_{\{k,l\}\in E^c}\left(GC_{kl}^tz_{kl}^t+GM_{kl}^t\left(1-\sum_{\tau=1}^{t-1}z_{kl}^\tau\right)\right)+\eta^t\\ =B^t+\left(\rho^{t-1}\eta^{t-1}\right),\quad t=2,\dots,T-1 \quad (29)$$

$$\sum_{k\in H^o}\left(FO_k^Ty_k^T+FM_k^1\sum_{\tau=1}^{T}y_k^\tau\right)+\sum_{k\in H^c}FM_k^1\left(1-\sum_{\tau=1}^{T-1}y_k^\tau\right)+\\ \sum_{\{k,l\}\in E^o}\left(GO_{kl}^Tz_{kl}^T+GM_{kl}^T\sum_{\tau=1}^{T}z_{kl}^\tau\right)+\sum_{\{k,l\}\in E^c}GM_{kl}^T\left(1-\sum_{\tau=1}^{t-1}z_{kl}^\tau\right)+\eta^T\\ =B^T+\left(\rho^{T-1}\eta^{T-1}\right) \quad (30)$$

10. Variable fixing

The meaning of the decision variables is only complete with the following constraints:

$$x_{ijkl}^t=0,\quad i,j\in N,\, i\neq j,\quad k,l\in N\setminus H,\, k\neq l,\quad t\in\mathcal{T} \quad (31)$$

$$u_{ijk}^t=0,\quad i,j\in N,\, i\neq j,\quad k\in N\setminus H, k\neq i,j,\quad t\in\mathcal{T} \quad (32)$$

$$v_{ijl}^t=0,\quad i,j\in N,\, i\neq j,\quad l\in N\setminus H, l\neq i,j,\quad t\in\mathcal{T} \quad (33)$$

$$s_{ij}^t=0,\quad i,j\in N\setminus H,\quad t\in\mathcal{T} \quad (34)$$

Accordingly, a MILP formulation for MPHLP, which we denote by $\mathcal{P}$ is the following:

$$
\begin{aligned}
& min \sum_{t\in\mathcal{T}}\sum_{i}\sum_{j>i}\left(W_{ij}^t+W_{ji}^t\right)\left(\sum_{k}\sum_{l\neq k}\alpha^t C_{kl}^t x_{ijkl}^t + C_{ij}^t s_{ij}^t + \sum_{k\neq i,j} C_{ik}^t u_{ijk}^t + \sum_{l\neq i,j} C_{lj}^t v_{ijl}^t\right) \\
& \quad + \sum_{t\in\mathcal{T}}\sum_{k\in H} F_k^t y_k^t + \sum_{t\in\mathcal{T}}\sum_{\{k,l\}\in E} G_{kl}^t z_{kl}^t \qquad (35) \\
& s.\,t.\; (1)-(34) \qquad (36) \\
& \quad x_{ijkl}^t \geq 0 \qquad i,j,k,l\in N,\ i\neq j,\ k\neq l,\ t\in\mathcal{T} \qquad (37) \\
& \quad u_{ijk}^t \geq 0 \qquad i,j,k\in N,\ i\neq j,\ k\neq i,j,\ t\in\mathcal{T} \qquad (38) \\
& \quad v_{ijl}^t \geq 0 \qquad i,j,l\in N,\ i\neq j,\ l\neq i,j,\ t\in\mathcal{T} \qquad (39) \\
& \quad s_{ij}^t \geq 0 \qquad i,j\in N,\ i\neq j,\ t\in\mathcal{T} \qquad (40) \\
& \quad y_k^t \in \{0,1\} \qquad k\in H \qquad (41) \\
& \quad z_{kl}^t \in \{0,1\} \qquad \{k,l\}\in E \qquad (42)
\end{aligned}
$$

Remark 1 *The model proposed above can easily accommodate the situation in which the set of demand nodes varies over the planning horizon. This can be achieved by simply setting to zero the flow originated and destined to a node in the periods in which the node is not 'active'.*

The MPHLP has the classical uncapacitated multiple allocation hub location problem as a particular case and thus it is NP-Hard.

The previous formulation is very heavy as was confirmed by a set of preliminary computational tests performed with the general solver CPLEX showing that the model above can only be solved with this solver for very small instances. For this reason we propose in the next section a heuristic procedure aiming at obtaining feasible solutions to the problem in an acceptable CPU time.

3 A local search-based procedure

In this section we present a local search-based procedure for obtaining feasible solutions to MPHLP.

As defined above, two types of decisions exist in the MPHLP, namely those related with the hub level network and those related with the flows, which in turn, determine the non-hub level network. One of the difficulties in solving the MPHLP arises from the need to integrate these two types of decisions. As we show next, if we knew the optimal hub level network, finding the optimal flows would be an easy problem to solve. Therefore we can simply follow widely known methodologies for uncapacitated multiple-allocation hub location problems (e.g. p-hub median problem). This situation motivates a procedure in which the focus is put in the hub level network. In particular, the procedure attempts to progressively find better feasible solutions by performing changes in the hub level network.

It is important to note that the knowledge about the operating hub edges in some period is sufficient to get the knowledge about the entire hub level network in that period. Accordingly, we can fully describe the hub level network from the set of operating hub edges because a decision made for some hub edge may determine decisions for the hubs. For instance, if a hub edge $\{k, l\}$ is established in the beginning of a time period t, this means that the hubs k and l must be active at that time. If this is not the case, the hubs should also be established. Conversely, if an existing hub edge $\{k, l\}$ finishes its operation at the end of period t hubs k and/or l may be removed if any of them becomes disconnected from the hub level network.

Starting with a feasible hub level network we propose a local search procedure performed in the hub level network. Each time a new hub level network is obtained, the corresponding non-hub network is determined as shown below and the cost of the corresponding feasible solution is evaluated.

Before going into the details of the heuristic procedure, we address several important aspects crucial for the procedure we propose.

3.1 Defining an initial feasible solution

We start by noting that we have a feasible hub level network as long as i) the hubs and hub edges make a connected sub-network, ii) the end points of the hub edges are operating hubs and iii) the budget constraints are not violated, we are facing a feasible hub level network.

A trivial feasible hub level network for MPHLP is obtained by taking the initial configuration and simply not changing it during the planning horizon. In such situation, the existing hubs and hub edges will be operating in all periods of the planning horizon and no new hub or hub edge will be installed. Such design leads to a feasible solution. The budget constraints are not violated because no network component is removed or installed throughout the planning horizon and it was assumed that the available budget is enough to maintain the initial configuration throughout the planning horizon.

Although the hub level network just described may lead to a poor feasible solution, it leads certainly to a solution very easy to obtain.

Once the network design is defined for all periods of the planning horizon, finding the optimal flows as well as the cost of the overall solution can be done easily as we show next.

3.2 Finding the optimal flows for a particular multi-period hub level network

Once the hub level network operating in each period in known, finding the optimal flows and thus the non-hub level network is something that can be easily done.

However, taking into account that no capacity constraints are considered in MPHLP, once the hub level network is defined in each period, an optimal way for routing the flow between each pair origin-destination consists simply in sending all the flow through the path with the smallest cost that connects the origin and the destination. This reasoning has been proposed in the literature for other multiple-allocation hub location problems. This was firstly stated by Campbell [9] and later on explored by Ebery et al. [21] and Ernst and Krishnamoorthy [23] among others.

For a particular time period $t \in \mathcal{T}$, we can build an auxiliary cost matrix D^t, which contains the costs between the nodes that can be directly connected (considering the hub level network in this period). For the infeasible connections we set the cost equal to a big M. This matrix is symmetric because the original matrix $[C^t_{ij}]_{i,j \in N}$ is also symmetric.

More formally, defining by H^t and E^t the set of hubs and hub edges, respectively, operating in period $t \in \mathcal{T}$, algorithm 1 can be used to determine matrix D^t.

Once the matrices D^t, $t \in \mathcal{T}$ are found, finding the non-hub level network can be easily done according to algorithm 2, where S^t denotes the set of non-hub edges in period $t \in \mathcal{T}$.

3.3 Computing the cost of a feasible solution

Once the network design is known for all periods of the planning horizon feasible values for the variables y and z are automatically defined. Denote by $\{\hat{z}^t_{kl}\}$ and $\{\hat{y}^t_k\}$ these values. If we also know the corresponding optimal flows, computing the cost of a feasible solution can be done straightforwardly according to algorithm 3. In this algorithm, P^t_{ij} denotes the

Algorithm 1 Finding matrix D^t for each $t \in \mathcal{T}$.

for $i = 1, \ldots, n$ **do**
  for $j = i+1, \ldots, n$ **do**
    if $i \in H^t$ and $j \in H^t$ and $\{i,j\} \notin E^t$ **then**
      $d^t_{ij} = d^t_{ji} = M$
    else if $i \in H^t$ and $j \in H^t$ and $\{i,j\} \in E^t$ **then**
      $d^t_{ij} = d^t_{ji} = \alpha^t C^t_{ij}$
    else if $i \notin H^t$ and $j \notin H^t$ **then**
      $d^t_{ij} = d^t_{ji} = M$
    else if $i \notin H^t$ and $j \in H^t$ **then**
      $d^t_{ij} = d^t_{ji} = C^t_{ij}$
    else if $i \in H^t$ and $j \notin H^t$ **then**
      $d^t_{ij} = d^t_{ji} = C^t_{ij}$
    end if
  end for
end for

set of hub edges in the shortest path from i to j in period t computed by using matrix D^t ($i, j \in H^t$, $t \in \mathcal{T}$) and *cost* denotes the cost of the actual feasible solution. As before, S^t denotes the set of non-hub edges in period $t \in \mathcal{T}$.

3.4 Local search procedure

A key element in the procedure we propose for the MPHLP is a neighborhood structure.

As we mentioned above, the knowledge of the hub edges operating in each period fully characterize the hub level network. Accordingly, for each feasible design for the hub level network over the planning horizon we define a neighborhood considering only the operating hub edges. In particular, we consider the variables z^t_{kl} introduced in section 2 which characterize the time period (if some) in which a hub edge $\{k,l\}$ is removed or established depending on whether the edge is in E^c or in E^o, respectively.

Consider a feasible design for the hub level network defined by the values $\{\hat{z}^t_{kl}\}$ of the z-variables. We define as the neighborhood of $\{\hat{z}^t_{kl}\}$ every solution $\{\overline{z}^t_{kl}\}$ obtained from the former as depicted in the next page. The main components of this exchanging mechanism regards steps 4a) b) and c). Move 4a) makes the closing period of some (existing) hub edge to be shifted from its current value t to another period t'; in move 4b), the opening of some

Algorithm 2 Finding the allocation of non-hubs to hubs for each $t \in \mathcal{T}$.

```
S^t ← ∅
for i = 1, ..., n do
  for j = i + 1, ..., n do
    if i ∉ H^t or j ∉ H^t then
      if W^t_ij > 0 or W^t_ji > 0 then
        Find the shortest path P between nodes i and j using matrix D^t
        if P has only one edge then
          S^t ← S^t ∪ {{i, j}}
        else
          Let (i, h_i) and (h_j, j) be the first and last edges of P, respectively.
          if i ∉ H^t and j ∉ H^t then
            S^t ← S^t ∪ {{i, h_i}, {h_j, j}}
          else if i ∉ H^t and j ∈ H^t then
            S^t ← S^t ∪ {{i, h_i}}
          else if i ∈ H^t and j ∉ H^t then
            S^t ← S^t ∪ {{h_j, j}}
          end if
        end if
      end if
    end if
  end for
end for
```

previously selected (new) hub edge is shifted from its current period t to another period t'; finally, move 4c), forces a hub edge that is not operaing in any period of the planning horizon to start operating in some period t'.

Algorithm 3 Evaluating the cost of a feasible solution.

```
cost ← ∑_{t∈T} ∑_{k∈H} F^t_k ŷ^t_k + ∑_{t∈T} ∑_{{k,j}∈E} G^t_{kl} ẑ^t_{kl}
for t = 1,...,T do
  for i = 1,...,n do
    for j = i+1,...,n do
      if i ∈ H^t ∧ j ∈ H^t then
        cost ← cost + (W^t_{ij} + W^t_{ji}) ∑_{{k,l}∈P^t_{ij}} α^t C^t_{kl}
      else if i ∈ H^t ∧ j ∉ H^t then
        Consider p ∈ H^t such that {j,p} ∈ S^t or {p,j} ∈ S^t
        cost ← cost + (W^t_{ij} + W^t_{ji}) (C^t_{jp} + ∑_{{k,l}∈P^t_{ip}} α^t C^t_{kl})
      else if i ∉ H^t ∧ j ∈ H^t then
        Consider p ∈ H^t such that {i,p} ∈ S^t or {p,i} ∈ S^t
        cost ← cost + (W^t_{ij} + W^t_{ji}) (C^t_{ip} + ∑_{{k,l}∈P^t_{pj}} α^t C^t_{kl})
      else
        Consider p ∈ H^t such that {i,p} ∈ S^t or {p,i} ∈ S^t
        Consider q ∈ H^t such that {j,q} ∈ S^t or {q,j} ∈ S^t
        cost ← cost + (W^t_{ij} + W^t_{ji}) (C^t_{ip} + C^t_{jq} + ∑_{{k,l}∈P^t_{pq}} α^t C^t_{kl})
      end if
    end for
  end for
end for
```

Neighborhood structure

1. Set $\overline{z}^t_{kl} \leftarrow \hat{z}^t_{kl} \quad (t \in \mathcal{T};\ \{k,l\} \in E)$
2. Select an hub edge $\{k', l'\} \in E$
3. Select a time period $t' \in \mathcal{T}$
4. Apply one of the following exchange moves depending on the situation
   (a) If $\{k', l'\} \in E^c$ let t be the period such that $\overline{z}^t_{k'l'} = 1$
   If $t \neq t'$ then $\overline{z}^t_{k'l'} = 0$ and $\overline{z}^{t'}_{k'l'} = 1$
   (b) If $\{k', l'\} \in E^o$ and exists $t \in \mathcal{T}$ such that $\overline{z}^t_{k'l'} = 1$ then
   If $t \neq t'$ then $\overline{z}^{t'}_{k'l'} = 1$, $\overline{z}^t_{k'l'} = 0$ else $\overline{z}^{t'}_{k'l'} = 0$
   (c) If $\{k', l'\} \in E^o$ and $\overline{z}^t_{k'l'} = 0\ \forall t \in \mathcal{T}$ then $\overline{z}^{t'}_{k'l'} = 1$

The previous instructions that together lead to a new set of values for the decision variables associated with the operating hub edges in each period of the planning horizon does not necessarily lead to a feasible hub level design. Note that feasibility is only achieved by hub networks that are connected in all periods of the planning horizon and also when the design does not violate the budget constraints.

When finding neighbors of a feasible hub level network we discard moves leading to disconnected hub level networks in some period. The connectivity of the hub level network in each period t is tested considering the submatrix D^t defined by the hubs and checking if there is at least one path of finite length (cost) between each pair of hubs.

Once a new solution is obtained in terms of the hub edges operating in each time period, the remainder components of the solution can be easily obtained, namely the hubs operating in each period and the non-hub level network (the latter by using Algorithm 2 after updating matrix D^t using Algorithm 1).

We have gathered so far the necessary ingredients to detail a heuristic procedure for the MPHLP.

Denote by $\mathcal{S}$ the set of feasible solutions to the problem and by $\mathcal{N}(.)$ the neighborhood of a feasible solution. The local search procedure is formalized in Algorithm 4.

Algorithm 4 Local search procedure.

Select a starting solution $s_0 \in \mathcal{S}$
repeat
  Select $\hat{s} \in \mathcal{N}(s_0) \cap \mathcal{S}$ such that $f(\hat{s}) = \min\limits_{s \in \mathcal{N}(s_0) \cap \mathcal{S}} \{f(s)\}$
  if $f(\hat{s}) < f(s_0)$ **then**
    Replace s_0 by $\hat{s}$
  end if
until $f(\hat{s}) \geq f(s_0)$
return s_0 // s_0 is the final feasible solution

4 Computational results

In this section we report the results of the computational tests performed in order to evaluate the possibility of solving the problem to optimality using the model proposed in section 2 and also in order to evaluate the performance of the methodology described in section 3. We start by describing the test instances considered. Afterwards the results are presented and analyzed.

4.1 Test instances

In order to perform a set of computational tests, 2 different classes of instances were generated: Class 1 containing pure randomly generated instances and Class 2 containing instances generated from the well-known AP data set (Ernst and Krishnamoorthy [22]).

In each class, the characteristics of the instances generated are the following:

- $n \in \{10, 15, 20, \ldots, 95, 100\}$.
- $T \in \{3, 6, 9, 12\}$.
- $|E^c| \in \{1, 2, 3\}$.
- $\alpha \in \{0, 7, 0.8, 0.9\}$.

For each combination of the above parameters, 3 instances were generated in order to get diversity. Combining the possibilities above we conclude that in total 4104 instances were generated - 2052 in each class.

The first step in the generation of one instance was to obtain the coordinates of the nodes. For the instances in Class 2 this is done as usual for the AP instances. In the case of instances in Class 1 the nodes were randomly generated in a 100×100 square. Once the coordinates of the nodes were available, the other parameters defining a single instance were generated as follows. In some situations a distinction is made between the instances in the two classes.

- For the instances in Class 1 the first period flows were generated according to $W_{ij}^1 \sim U\{10, \ldots, 20\}$.

  In the case of the AP instances (Class 2), for the first period we considered the flows available in the literature.

  Independently from the instance class, for $t = 2, \ldots, T$, $W_{ij}^t = \varphi \times W_{ij}^{t-1}$ with $\varphi \sim U[1.05; 1.10]$.

The random factor φ was generated each time a new value is to be obtained. This comment is also valid for the parameters below.

- $FO_k^1 \sim U[500; 700]$, $k \in H^o$.

  For $t = 2, \ldots, T$ and $k \in H^o$, $FO_k^t = \varphi \times FO_k^{t-1}$ with $\varphi \sim U[1.05; 1.10]$.

- $FC_k^1 \sim U[200; 300]$, $k \in H^c$.

  For $t = 2, \ldots, T$ and $k \in H^c$, $FC_k^t = \varphi \times FC_k^{t-1}$ with $\varphi \sim U[1.05; 1.10]$.

- $FM_k^1 \sim U[300; 400]$, $k \in H$.

  For $t = 2, \ldots, T$ and $k \in H$, $FM_k^t = \varphi \times FM_k^{t-1}$ with $\varphi \sim U[1.10; 1.20]$.

- $GO_{kl}^1 \sim U[120; 130]$, $\{k, l\} \in E^o$.

  For $t = 2, \ldots, T$ and $\{k, l\} \in E^o$, $GO_{kl}^t = \varphi \times GO_{kl}^{t-1}$ with $\varphi \sim U[1.05; 1.10]$.

- $GC_{kl}^1 \sim U[80; 85]$, $\{k, l\} \in E^c$.

  For $t = 2, \ldots, T$ and $\{k, l\} \in E^c$, $GC_{kl}^t = \varphi \times GC_{kl}^{t-1}$ with $\varphi \sim U[1.05; 1.10]$.

- $GM_{kl}^1 \sim U[100; 110]$, $\{k, l\} \in E$.

  For $t = 2, \ldots, T$ and $\{k, l\} \in E$, $GM_{kl}^t = \varphi \times GM_{kl}^{t-1}$ with $\varphi \sim U[1.10; 1.20]$.

- $\rho^t = 1.1$, $t \in \mathcal{T}$.

- For $t = 1, \ldots, T$, $B^t = \xi \times \Psi(t)$ with $\Psi(t) = \sum_{k \in H^c} FM_k^t + \sum_{\{k,l\} \in E^c} GM_{kl}^t$ and ξ defined as follows:

  $$\xi = \begin{cases} 3 & \text{if } t = 1 \text{ or } t = T \\ 1 + 0.2T - 0.2(t-1) & \text{if } t \in \{2, ..., T-1\} \end{cases}$$

  The automatic generation of values for the budget available in each period is tricky as one easily gets an infeasible instance or an instance that is not interesting (e.g. the available budget is too large leading to too loose budget constraints).

- The values C_{ij}^t were set equal to half of the euclidean distances between the corresponding pairs of nodes. Accordingly, these parameters, are considered static in our instances. Note that in the case of the AP instances, the distances are firstly divided by 1000 as usual in these instances.

Note that the costs FO_k^t, FC_k^t, FM_k^t, GO_{kl}^t, GC_{kl}^t, GM_{kl}^t as well as the factors ρ^t are generated similarly for both classes of instances.

Finally, the information about one instance is only complete when an initial configuration is defined. This was done as follows: One node was randomly chosen. From this node, the least cost edge was found and added to the configuration. Starting from the end node of this link, again, the least cost edge is found and added to the configuration. We proceed until we obtain the desired number of edges in the initial configuration.

4.2 Analysis of the results

In order to get a perception of the impact from using the methodology proposed in the previous section, not only did we run the heuristic but also we considered the feasible solution associated with the initial configuration if such configuration is never changed during the planning horizon. Note that the way the budget is generated in each period assures that the initial configuration leads always to a feasible solution to the problem which simply consists in keeping such configuration in all periods.

All computational tests were performed on a Machine with a Intel(R) Core(TM)2 Duo processor, 3.00 GHz, 3.21 GB of RAM.

An attempt was made to solve the problem to optimality using a general solver. We also attempted to solve the linear relaxation so that lower bounds could be obtained. Such attempt was made using IBM ILOG CPLEX 12.1. However, successful results were obtained only for instances with up to 15 nodes. For the instances with 20 nodes, the solver could only tackle instances with 3 times periods. This attempt shows the relevance of developing heuristic approaches for the problem we are addressing.

The results obtained with the heuristic presented in the previous section are summarized in Tables 1 and 2. In each table, the first column depicts the values of the parameters that were used for summarizing the results. In columns 2-4 we can observe the average, minimum and maximum percentage of improvement achieved with the solution found by the heuristic when compared with the solution induced by the initial configuration. The percentage of improvement was computed according to $(V_I - V_H)/V_I * 100$ where V_I denotes the cost of the feasible solution induced by the initial configuration whereas V_H denotes the cost of the solution provided by the heuristic. In columns 5-7 we can observe the cpu time (average, minimum and maximum) required by the heuristic to find a feasible configuration.

Table 1 presents the results for the instances in Class 1 whereas Table 2 contains the same information but for the instances in Class 2. Each table is divided into 3 sub-tables: the first organized by the different values of n; the second by the values of T, the third by the values of α. Accordingly, each row in Tables 1a and 2a corresponds to 108 instances; each row in Tables 1b and 2b corresponds to 513 instances; finally, each row in Tables 1c and 2c corresponds to 684 instances.

Observing Tables 1a and 2a we conclude the in terms of the improvement over the initial configuration, no pattern can be found. The instances seem to behave in a relative similar manner. It is interesting also to note that the improvement in the AP instances seem slightly larger than the improvement that occurred in the purely randomly generated instances. This is more evident when we observe the maximum improvement values that in the AP instances reached around 140% but in the instances of Class 1, did not go over 81%. As it was naturally expected, the cpu time increased roughly by a factor of 1000 when we moved from instances with 10 nodes to instances with 100 nodes.

Observing tables 1b and 2b we see that as far as the average improvement is concerned, the two classes behave differently. In the case of class 1, (1b) the average improvement increased with T which gives an indication that in instances with a longer planning horizon more time is given for improving the initial configuration. this behavior was not observed in the instances in Class 2. In this case, the instances in which a better (average) improvement was achieved are the ones corresponding to $T = 6$. Finally, as expected, the cpu time increases by a factor of roughly 30 when we go from $T = 3$ to $T = 12$.

As far as the value of the discount factor α is concerned, the results show no pattern worth mentioning. The average results were not significantly influenced by the value of α.

The results presented in the previous sub-section give an overall figure for the tested instances. Nevertheless, in order to get a better perception of the changes over time in the network structure we considered three instances in Class 2 and represent the network structure below. The three instances considered refer to one AP data instance with 25 nodes and 6 times periods. The only difference among the three instances is the value of

n	Improvement (%) average	minimum	maximum	Time (sec.) average	minimum	maximum
10	13.23	0.00	55.05	1.40	0.06	5.02
15	14.40	0,00	54.28	3.60	0.17	14.28
20	17.34	2.96	41.48	7.81	0.40	57.35
25	22.31	5,47	80.93	9.34	0.71	208.31
30	20.30	0.28	54.49	23.37	1.19	127.33
35	19.09	4,75	43.17	32.65	2.34	143.48
40	19.49	4.08	41.47	65.20	3.05	660.62
45	19.08	5,49	62.29	70.97	4.44	250.37
50	16.14	4.08	36.83	102.13	6.65	660.74
55	16.50	5,35	39.43	121.58	7.99	461.81
60	16.51	2.78	42.53	143.64	10.48	663.85
65	17.86	8,70	35.41	205.76	15.56	907.86
70	13.84	4.97	33.82	281.67	18.89	1535.58
75	13.26	4,63	35.84	325.22	22.29	1399.72
80	14.25	3.61	51.48	441.46	26.12	2227.87
85	15.78	5,21	42.38	498.49	35.50	2097.63
90	16.26	6.60	39.24	679.61	37.24	3476.77
95	13.81	3,01	41.42	803.33	47.00	4587.93
100	15.54	7.39	30.81	973.98	65.25	13198.27

(a) Results summarized by the different values of the number of nodes - n.

T	Improvement (%) average	minimum	maximum	Time (sec.) average	minimum	maximum
3	11.93	0.00	46.23	26.32	0.06	166.49
6	14.38	0.00	55.05	113.25	0.21	1344.79
9	19.55	3.01	61.38	276.50	0.77	3144.38
12	20.47	0.41	80.93	593.92	1.01	13198.27

(b) Results summarized by the different values of the number of periods - T.

α	Improvement (%) average	minimum	maximum	Time (sec.) average	minimum	maximum
0.7	16.72	0.00	66.39	245.50	0.06	3476.77
0.8	16.53	0.00	64.93	268.03	0.06	13198.27
0.9	16.49	0.00	80.93	243.96	0.06	4587.93

(c) Results summarized by the different values of the discount factor - α.

Table 1: Results for the instances in Class 1.

	Improvement (%)			Time (sec.)		
n	average	minimum	maximum	average	minimum	maximum
10	15.02	0.34	52.41	1.00	0.05	6.29
15	20.14	0.44	61.19	4.34	0.17	23.84
20	18.06	1.83	40.21	10.66	0.39	49.80
25	29.03	2.80	87.33	16.49	0.82	125.61
30	20.12	1.14	60.44	27.85	1.25	241.96
35	16.29	2.82	57.72	69.75	2.19	254.68
40	19.61	0.98	84.81	91.10	3.14	530.54
45	26.96	1.40	79.38	150.29	4.81	579.11
50	21.92	1.40	85.96	159.20	6.84	731.91
55	32.81	0.18	139.32	309.89	9.11	2083.19
60	24.83	2.08	118.11	300.71	12.19	1672.34
65	28.78	3.93	126.71	556.78	16.12	2868.55
70	28.08	2.40	114.30	617.91	20.09	3255.08
75	21.51	1.78	104.87	639.57	26.52	3001.93
80	23.49	2.42	78.95	771.72	27.37	4092.75
85	22.71	1.66	78.43	887.22	42.54	3747.88
90	17.98	1.77	72.38	1167.61	37.65	5932.95
95	27.86	0.90	82.01	1700.23	50.29	6364.38
100	19.90	4.02	66.26	1737.13	58.01	6652.66

(a) Results summarized by the different values of the number of nodes - n.

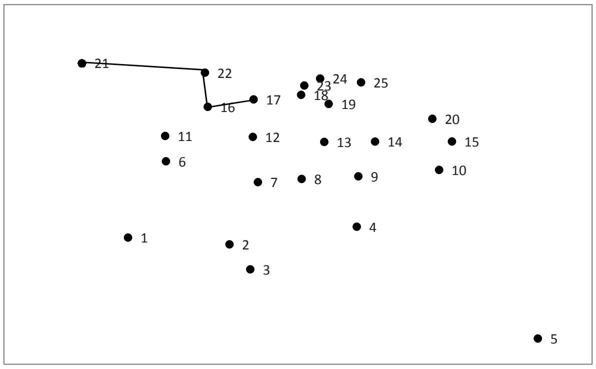
	Improvement (%)			Time (sec.)		
T	average	minimum	maximum	average	minimum	maximum
3	23.12	3.13	118.11	38.16	0.05	294.89
6	27.06	2.70	139.32	230.45	0.20	2399.22
9	21.57	2.40	104.87	578.01	0.97	4185.16
12	19.85	0.18	107.94	1096.63	0.73	6652.66

(b) Results summarized by the different values of the number of periods - T.

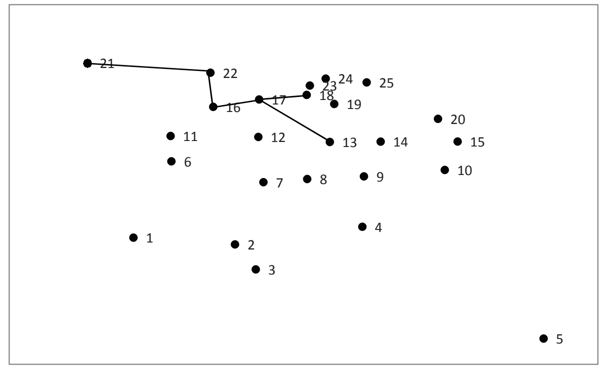
	Improvement (%)			Time (sec.)		
α	average	minimum	maximum	average	minimum	maximum
0.7	22.97	0.23	139.32	483.91	0.05	6364.38
0.8	22.91	0.20	137.97	487.10	0.06	6572.16
0.9	22.82	0.18	136.91	486.43	0.05	6652.66

(c) Results summarized by the different values of the discount factor - α.

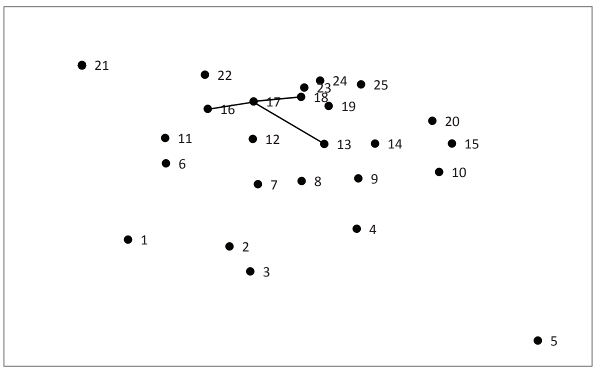
Table 2: Results for the instances in Class 2.



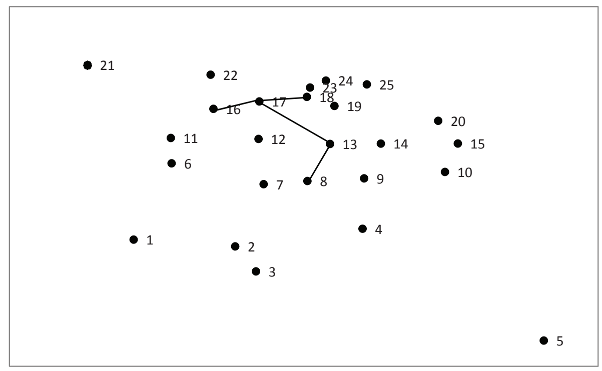
(a) Initial configuration.



(b) Period 1.



(c) Periods 2 and 3.



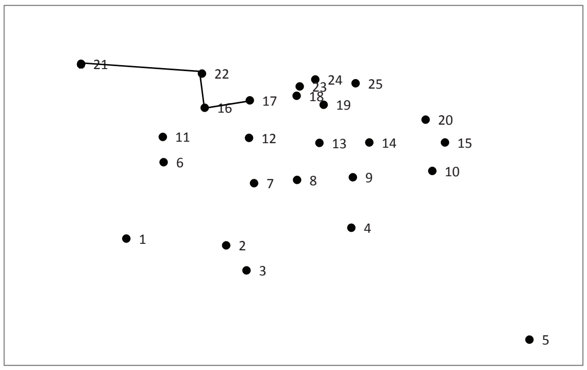
(d) Periods 4, 5 and 6.

Figure 3: Hub level network for one of the AP data instances with 25 nodes and 6 periods - $\alpha = 0.7$.

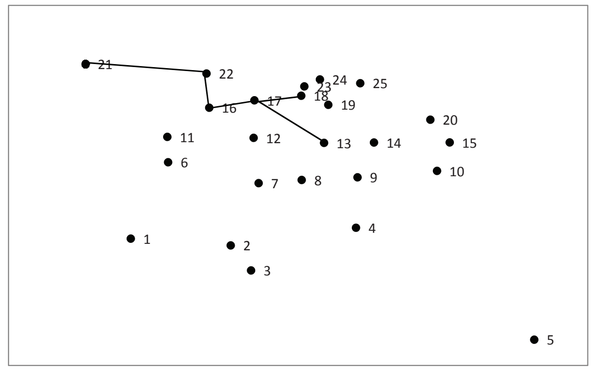
α. In particular we consider the three values of α that were mentioned above: 0.7, 0.8 and 0.9. The results can be observed in Figures 3 and 4.

In these figures apart from the initial configuration we present the different configurations that the network assume over time.

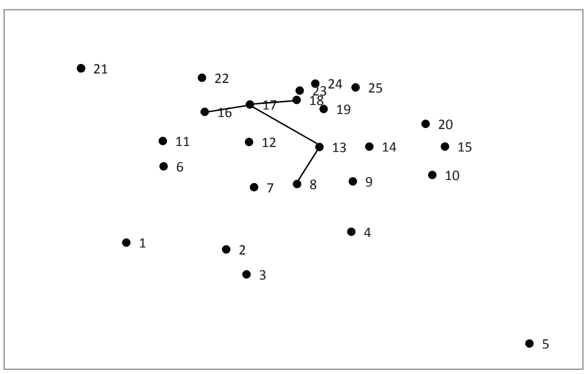
In the examples considered, we can observe that an initially peripheral network was progressively changed to a more central structure. It should also be noted that a slight change in the value of α led to different evolutions of the network structure. This example



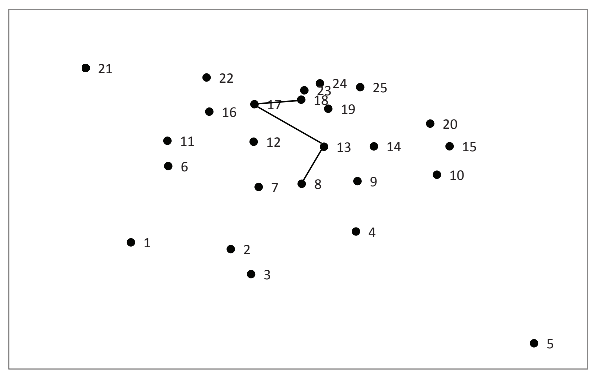
(a) Initial configuration.



(b) Period 1.



(c) Period 2.



(d) Periods 3, 4, 5 and 6.

Figure 4: Hub level network for one of the AP data instances with 25 nodes and 6 periods - $\alpha = 0.8, 0.9$.

is illustrative of the impact that the value of the discount factor may have in a hub level network. In terms of the total cost we have:

For $\alpha = 0.7$, the initial configuration induces a feasible solution having cost 601740.99. The multi-period solution which network structure is represented in figures 3b-3d has a cost of 455597.58. This difference represents an improvement of 24.28%.

For $\alpha = 0.8$, the cost of the feasible solution induced by the initial configuration is 604024.29 whereas the cost of the solution found by the heuristic (which network structure is depicted in Figures 4b-4d) is 451015.05. The difference represents an improvement of 25.33%.

Finally, for $\alpha = 0.9$, the costs of the initial feasible solution is 605997.97 and the heuristic reached a feasible solution with a cost of 454468.79. The improvement percentage was 25.00%.

5 Conclusion

In this paper, a multi-period phase-in/phase-out hub location problem was considered in which several classical assumptions often considered in hub location literature were relaxed. A mathematical programming formulation was developed. Due to the complexity of the problem, a heuristic approach was also proposed which is a local search based procedure. The results show that by using the proposed approach, significant improvements were achieved if we compare the solutions obtained with the solution induced by a static network structure.

This work raises several questions which require further research. The first regards the routing costs. The model we considered in this paper assume that there is a discount factor for the flow that traverses the hub edges. However no mechanism was considered for assuring that the flow in the hub edges exceeds the flow in the non hub edges and thus that the discount factor is consistent. This is a well-known weakness in many hub location problems. Nevertheless it deserves more research, namely in the same context that we considered in this paper.

Another important research direction regards the development of lower bounds for the problem at hand. This is important as a means for properly evaluating the quality of feasible solutions. Currently, the quality of the solutions produced by the methodology that we developed was measured only by comparing those solutions with the ones induced by the network structure that is operating at the beginning of the planning horizon.

Acknowledgement

The authors would like to express their gratitude to the anonymous referees whose comments helped to improved the paper. This work is supported by National Funding from FCT - Fundação para a Ciência e a Tecnologia, under the projects PEst-OE/MAT/UI0152 (CIO/FCUL) and PEst-OE/MAT/UIO297/2011 (CMA/FCT/UNL).

References

[1] S. Abdinnour-Helm. A hybrid heuristic for the uncapacitated hub location problem. *European Journal of Operational Research*, 106:489–499, 1998.

[2] S. Abdinnour-Helm and M. A. Venkataramanan. Solution approaches to hub location problems. *Annals of Operations Research*, 78:31–50, 1998.

[3] M. Albareda-Sambola, E. Fernández, Y. Hinojosa, and J. Puerto. The multi-period incremental service facility location problem. *Computers & Operations Research*, 36: 1356–1355, 2008.

[4] S. Alumur and B. Y. Kara. Network hub location problems: the state of the art. *European Journal of Operational Research*, 190:1–21, 2008.

[5] S. Alumur and B. Y. Kara. The design of single allocation incomplete hub networks. *Transportation Research Part B*, 43:936–951, 2009.

[6] S. Alumur and B. Y. Kara. A hub covering network design problem for cargo applications in Turkey. *Journal of the Operational Research Society*, 60:1349–1359, 2009.

[7] N. Boland, M. Krishnamoorthy, A. T. Ernst, and J. Ebery. Preprocessing and cutting for multiple allocation hub location problems. *European Journal of Operational Research*, 155:638–653, 2004.

[8] J. Campbell. Locating transportation terminals to serve an expanding demand. *Transportation Research Part B: Methodological*, 24:173–192, 1990.

[9] J. F. Campbell. Integer programming formulations of discrete hub location problems. *European Journal of Operational Research*, 72:387–405, 1994.

[10] J. F. Campbell. Hub location and the p-hub median problem. *Operations Research*, 44:1–13, 1996.

[11] J. F. Campbell, A. T. Ernst, and M. Krishnamoorthy. Hub location problems. In Z. Drezner and H. W. Hamacher, editors, *Facility Location: Applications and Theory*, pages 373–407. Springer, 2002.

[12] J. F. Campbell, A. T. Ernst, and M. Krishnamoorthy. Hub arc location problems: Part I - Introduction and Results. *Management Science*, 51:1540–1555, 2005.

[13] J. F. Campbell, A. T. Ernst, and M. Krishnamoorthy. Hub arc location problems: Part II - Formulations and Optimal Algorithms. *Management Science*, 51:1556–1571, 2005.

[14] J.-F. Chen. A hybrid heuristic for the uncapacitated single allocation hub location problem. 35:211–220, 2007.

[15] I. Contreras, J.-F. Cordeau, and G. Laporte. The dynamic uncapacitated hub location problem. Technical Report 41, CIRRELT, 2009.

[16] I. Contreras, J. Díaz, and E. Fernández. Lagrangean relaxation for the capacitated hub location problem with single assignment. *OR Spectrum*, 31:483–505, 2009.

[17] I. Contreras, E. Fernández, and A. Marín. Tight bounds from a path based formulation for the tree of hub location problem. *Computers & Operations Research*, 36:3117–3127, 2009.

[18] I. Contreras, E. Fernández, and A. Marín. The tree of hubs location problem. *European Journal of Operational Research*, 202:390–400, 2010.

[19] I. Correia, S. Nickel, and F. Saldanha da Gama. Single-assignment hub location problem with multiple capacity levels. *Transportation Research Part B*, 44:1047–1066, 2010.

[20] I. Correia, S. Nickel, and F. Saldanha-da-Gama. The capacitated single-allocation hub location problem revisited: A note on a classical formulation. *European Journal of Operational Research*, 207:92–96, 2010.

[21] J. Ebery, M. Krishnamoorthy, A. Ernst, and N. Boland. The capacitated multiple allocation hub location problem: Formulations and algorithms. *European Journal of Operational Research*, 120:614–631, 2000.

[22] A. T. Ernst and M. Krishnamoorthy. Efficient algorithms for the uncapacitated single allocation p-hub median problem. *Location Science*, 4:139–154, 1996.

[23] A. T. Ernst and M. Krishnamoorthy. Exact and heuristic algorithms for the uncapacitated multiple allocation p-hub median problem. *European Journal of Operational Research*, 104:100–112, 1998.

[24] A. T. Ernst and M. Krishnamoorthy. Solution algorithms for the capacitated single allocation hub location problem. *Annals of Operations Research*, 86:141–159, 1999.

[25] S. Gelareh. *Hub location models in public transport planning*. PhD thesis, Tecnical University of Kaiserslautern, Germany, 2008.

[26] S. Gelareh and S. Nickel. New approaches to hub location problems in public transport planning. Technical Report 133, Fraunhofer Institut für Techno- und Wirtschaftsmathematik (ITWM), Kaiserslautern, Germany, 2008. Available at www.itwm.fhg.de.

[27] B. Y. Kara and M. R. Taner. Hub location problems: The location of interacting facilities. In H. A. Eiselt and V. Marianov, editors, *Foundations of location analysis*, pages 273–288. Springer, 2011.

[28] J. Klincewicz. Heuristics for the p-hub location problem. *European Journal of Operational Research*, 53:25–37, 1991.

[29] J. Klincewicz. Avoiding local optima in the p-hub location problem using tabu search and grasp. *Annals of Operations Research*, 40:283–302, 1992.

[30] M. Labbé, H. Yaman, and E. Gourdin. A branch and cut algorithm for the hub location problems with single assignment. *Mathematical Programming*, 102:371–405, 2005.

[31] G. Mayer and B. Wagner. Hublocator: an exact solution method for the multiple allocation hub location problem. *Computers & Operations Research*, 29:715–739, 2002.

[32] M.T. Melo, S. Nickel, and F. Saldanha-da-Gama. Dynamic multi-commodity capacitated facility location: A mathematical modeling framework for strategic supply chain planning. *Computers & Operations Research*, 33:181–208, 2006.

[33] S. Nickel, A. Schobel, and T. Sonneborn. Hub location problems in urban traffic networks. In J. Niittymaki and M. Pursula, editors, *Mathematics Methods and Optimization in Transportation Systems*, pages 1–12. Kluwer Academic Publishers, 2001.

[34] M. O'Kelly. A quadratic integer problem for the location of interacting hub facilities. *European Journal of Operational Research*, 32:393–404, 1987.

[35] M. O'Kelly. The location of interacting hub facilities. *Transportation Science*, 20: 92–106, 1989.

[36] D. Skorin-Kapov, J. Skorin-Kapov, and M. O'Kelly. On tabu search for the location of interacting hub facilities. *European Journal of Operational Research*, 73:502–509, 1994.

[37] D. Skorin-Kapov, J. Skorin-Kapov, and M. O'Kelly. Tight linear programming relaxations of uncapacitated p-hub median problems. *European Journal of Operational Research*, 94:582–593, 1996.

[38] H. Topcuoglu, F. Corut, M. Ermis, and G. Yilmaz. Solving the uncapacitated hub location problem using genetic algorithms. *Computers & Operations Research*, 32: 967–984, 2005.

[39] H. Yaman. Star p-hub median problem with modular arc capacities. *Computers & Operations Research*, 35:3009–3019, 2008.

[40] H. Yaman and G. Carello. Solving the hub location problem with modular link capacities. *Computers & Operations Research*, 32:3227–3245, 2005.