



A capacitated hub location problem under hose demand uncertainty



Merve Meraklı^{a,b,*}, Hande Yaman^a

^a Department of Industrial Engineering, Bilkent University, Ankara, Turkey

^b Department of Integrated Systems Engineering, The Ohio State University, Columbus, OH, USA

ARTICLE INFO

Article history:

Received 8 December 2016

Revised 10 May 2017

Accepted 14 June 2017

Available online 20 June 2017

Keywords:

Hub location

Multiple allocation

Capacitated hubs

Demand uncertainty

Robustness

Hose model

Benders decomposition

ABSTRACT

In this study, we consider a capacitated multiple allocation hub location problem with hose demand uncertainty. Since the routing cost is a function of demand and capacity constraints are imposed on hubs, demand uncertainty has an impact on both the total cost and the feasibility of the solutions. We present a mathematical formulation of the problem and devise two different Benders decomposition algorithms. We develop an algorithm to solve the dual subproblem using complementary slackness. In our computational experiments, we test the efficiency of our approaches and we analyze the effects of uncertainty. The results show that we obtain robust solutions with significant cost savings by incorporating uncertainty into our problem.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

Hubs are used commonly in many-to-many distribution systems that arise in transportation and telecommunications applications. Flows from many origins to many destinations are consolidated at hubs and routed together to benefit from economies of scale. Many variants of hub location problems have been studied in the last few decades. Given a set of nodes with pairwise traffic demands, the hub location problem decides on the locations of the hubs and the routes of traffic demands to minimize some performance measure. This measure can be related with the system cost or the quality of service. The system cost includes the cost of routing the traffic in the hub network and it may include the fixed cost of locating hubs if the number of hubs is not fixed. In some variants, direct shipments between nonhub nodes are allowed, in others all the traffic is routed through at least one hub. Also, there are variants of the problem where a nonhub node can send and receive traffic through multiple hubs and others where there is a restriction on the number of hubs that a nonhub node can use. The first setting is known as the multiple allocation setting. In this paper, we study a hub location problem with multiple allocation, fixed costs for installing capacitated hubs and no direct shipments.

Most studies in the hub location literature are based on the assumption that the pairwise demands are known with certainty. However, this is very difficult to justify in practice since strategic decisions such as hub location decisions are often taken before observing the actual demand and the demand fluctuates over time. In this study, we incorporate the demand uncertainty into the capacitated multiple allocation hub location problem. In this setting, demand uncertainty affects both the feasibility of a hub network and its associated cost. To hedge against demand uncertainty, we use a robust optimization framework: among all hub networks that are feasible for all possible demand realizations, we would like to find one that minimizes the worst case total cost (for more on robust optimization see, e.g., [Atamtürk \(2006\)](#); [Ben-Tal et al. \(2004\)](#); [Ben-Tal and Nemirovski \(1998\)](#); [1999](#); [2008](#)); [Bertsimas and Sim \(2003\)](#); [2004](#)); [Mudchanatongsuk et al. \(2008\)](#); [Ordóñez and Zhao \(2007\)](#); [Yaman et al. \(2001\)](#); [2007](#))).

We represent the uncertainty with a special polyhedral uncertainty model known as the hose model. The parameters of this model are aggregate traffic upper bounds for each node. Any non-negative demand vector in which the sum of traffic demands that each node can send and receive does not exceed the traffic upper bound for that node is a possible demand realization. The hose model was proposed by [Duffield et al. \(1999\)](#) and [Fingerhut et al. \(1997\)](#) to design virtual private networks. It has several advantages compared to other uncertainty models: it asks to estimate a parameter for each node rather than for each pair of nodes. This aggregation reduces the statistical variability and errors. It has resource-sharing flexibility and is not a conservative

* Corresponding author at: Department of Integrated Systems Engineering, The Ohio State University, Columbus, OH, USA.

E-mail addresses: merve.merakli@bilkent.edu.tr (M. Meraklı), hyaman@bilkent.edu.tr (H. Yaman).

model in which each origin-destination traffic demand can take its worst case value simultaneously. Due to these advantages, the hose model has been used as an uncertainty model in many studies following its introduction (some examples are Altın et al. (2007); 2011); Chekuri et al. (2007); Italiano et al. (2006)).

Recently, Meraklı and Yaman (2016) study the uncapacitated multiple allocation p -hub median problem with polyhedral demand uncertainty. They present a mixed integer programming model and apply Benders decomposition. Their results show that algorithms based on decomposition are very efficient compared to solving the model with an off-the-shelf solver. They also observe that it is possible to obtain significant cost savings by incorporating demand uncertainty into the problem. In the uncapacitated problem, the demand only affects the routing costs. In addition, it is known that when hub locations are given, each traffic demand is routed on a shortest path from its origin to its destination independently of the amount of demand. As a result, it is possible to hedge against uncertainty with minor changes in the network. These are not true when capacity constraints are imposed for hubs.

In this paper, we present a model for the capacitated hub location problem with multiple allocation and hose demand uncertainty. Our initial computational experiments showed that the model is much harder to solve compared to its deterministic counterpart. We propose two exact algorithms based on Benders reformulations and give an algorithm to solve the dual subproblem using complementary slackness. We test the efficiency of these algorithms using instances from the literature. We also perform experiments to investigate the changes in the hub locations and costs as a result of demand uncertainty. We observe that ignoring demand uncertainty may result in high routing costs and congested hubs. Unlike the observations for the uncapacitated problem, when capacity constraints are imposed, one may need to make major changes in the hub locations to hedge against uncertainty.

The rest of the paper is organized as follows. In Section 2, we review the related literature. In Section 3, we first present a nonlinear model and then derive a compact linear mixed integer programming model. We give two Benders reformulations in Section 4. We report the results of computational experiments in Section 5 and conclude the paper in Section 6.

2. Literature review

In the last few decades, hub location problems have received a lot of attention both in telecommunications and transportation literatures. Here we limit ourselves to related studies and refer the reader to surveys in Campbell (1994b), and Alumur and Kara (2008); Campbell et al. (2002); Campbell and O’Kelly (2012); Klinecicz (1998); O’Kelly and Miller (1994) and Farahani et al. (2013) for further information.

The multiple allocation hub location problem is first formulated by Campbell (1994a). Boland et al. (2004); Camargo et al. (2008); Cánovas et al. (2007); Ebery et al. (2000); Ernst and Krishnamoorthy (1998a); Hamacher et al. (2004); Klinecicz (1996); Marín (2005b); Mayer and Wagner (2002) and Contreras et al. (2011a) propose methods to solve this problem. The version of the problem where there is no cost for opening hubs but the number of hubs is fixed to p is first formulated by Campbell (1992). Alternative formulations are given by Campbell (1994a); Skorin-Kapov et al. (1996) and Ernst and Krishnamoorthy (1998a). Campbell (1996) and Ernst and Krishnamoorthy (1998a); 1998b) propose exact and heuristic solution algorithms.

Among the studies cited above, several propose Benders decomposition based approaches. Camargo et al. (2008) propose three different algorithms: the classical Benders decomposition approach, which adds a single cut at each iteration, a multi-cut version in which Benders cuts are generated for each origin-

destination pair and a variant which terminates when an ϵ -optimal solution is obtained. Contreras et al. (2011a) propose a Benders decomposition in which they generate cuts for each candidate hub location instead of each origin-destination pair. Camargo et al. (2009) propose two Benders decomposition algorithms to solve the variant of the problem where the cost is a piecewise-linear concave function. Gelareh and Nickel (2011) study a problem with an incomplete hub network and solve this problem with a Benders decomposition algorithm.

Capacitated variants of the hub location problems received less attention in the literature compared to the uncapacitated versions. The first mixed integer linear programming formulation for the capacitated multiple allocation hub location problem (CMAHLP) is proposed by Campbell (1992) using four indexed variables. Ebery et al. (2000) provide formulations with three indices and devise a heuristic algorithm to solve large instances. In order to strengthen these formulations, Boland et al. (2004) propose preprocessing procedures and valid inequalities, which lead to a significant reduction in the computation times. Marín (2005a) also provides new formulations and resolution techniques to obtain better computational results and succeeds to solve instances with up to 75 nodes. Sasaki and Fukushima (2003) consider a capacitated multiple allocation hub location problem where a capacity constraint is applied both on hubs and arcs and a flow can go through at most one hub on its way from origin to destination. They devise a branch and bound algorithm and perform computational studies on the CAB data set.

There are also Benders decomposition applications for the capacitated multiple allocation hub location problems. Rodríguez-Martín and Salazar-González (2008) consider a capacitated hub location problem with multiple allocation on an incomplete hub network. They provide a formulation and develop two exact solution algorithms. The first one utilizes classical Benders decomposition approach whereas the second employs a nested two level algorithm based on Benders decomposition. They show that the latter outperforms the classical Benders decomposition approach in terms of computation times. Contreras et al. (2012) also study a related capacitated hub location problem in which the capacities installed on each hub is not a parameter but a decision variable. They devise a Benders decomposition algorithm in which the subproblem is a transportation problem. They apply Pareto-optimal Benders cuts and reduction tests to improve the convergence of the algorithm.

The studies that incorporate data uncertainty into hub location problems is rather limited. Marianov and Serra (2003) study the problem in an air transportation network where hubs are $M/D/c$ queues and the probability that the number of planes in the queue exceeds a certain number is bounded above. This restriction is then reformulated as a capacity constraint for the hubs. The authors propose a tabu search based heuristic method to solve this problem. Yang (2009) decides on hub locations and flight routes under demand uncertainty using two-stage stochastic programming. The first stage involves the decision on the locations of the hubs to open. In the second stage, routes are determined after demand realizations are observed. Sim et al. (2009) incorporate service level considerations using chance constraints when travel times are normally distributed. They propose several heuristic algorithms. Contreras et al. (2011b) consider the uncapacitated multiple allocation hub location problem under demand and transportation cost uncertainty. They show that the stochastic models for this problem with uncertain demands or transportation costs dependent to a single uncertain parameter are equivalent to the deterministic problem with mean values. This is not the case for the problem with stochastic independent transportation costs. This latter problem is solved using Benders decomposition and a sample average scheme. They use the AP data set to test the

efficiency and effectiveness of the proposed models and algorithms. Alumur et al. (2012) consider uncertainty both in fixed costs and demands. They use a minimax regret approach and stochastic programming to hedge against uncertainty. Shahabi and Unnikrishnan (2014) propose mixed integer conic quadratic programming formulations for hub location problems with ellipsoidal demand uncertainty. Meraklı and Yaman (2016) study the uncapacitated multiple allocation p -hub median problem with hose demand uncertainty and present Benders decomposition based algorithms.

In this study, we incorporate both demand uncertainty and capacity constraints for hubs into the multiple allocation hub location problem. This results in a more challenging problem compared to the uncapacitated case since demands have an impact both in the cost and feasibility of a solution. The decomposition approaches also need further analysis to be effectively used. Our results show that it is even more critical to consider demand uncertainty in the case of the capacitated problem since the deterministic solution may not be feasible when the realized demand is different from the estimated one.

3. MIP formulation

In this section we formulate the robust CMAHLP under hose demand uncertainty. In this problem, nonhub nodes can be connected to multiple hubs and a capacity constraint on the incoming flow at each hub is imposed. The deterministic version of this problem has been formulated in several ways in the literature. We use the formulation proposed by Hamacher et al. (2004) as a starting point. This formulation is devised for the uncapacitated version of the problem, hence we adjust it by adding a set of capacity constraints as proposed in Ebery et al. (2000).

We are given a set N of demand points. Let $H \subseteq N$ be the set of possible hub locations and C be the set of commodities such that $C = \{(i, j) : i, j \in N, i \neq j\}$, i.e., any ordered pair of distinct nodes is a commodity. The demand from node i to node j is assumed to be known in the deterministic problem and is denoted by w_{ij} . We define the remaining problem parameters as follows: f_k is the fixed cost of opening a hub facility at node k , a_k is the capacity of the hub at node k , d_{ij} is the unit cost of transshipment from node i to node j and χ , α and δ are the cost multipliers of collection, transfer between hubs and distribution, respectively. The cost of sending one unit of flow from node i to node j through hubs k and m in this order is expressed as $c_{ijkm} = \chi d_{ik} + \alpha d_{km} + \delta d_{mj}$.

First we present the MIP formulation for the deterministic CMAHLP. The decision variables of this model are y_k , the binary variable taking value of 1 if there is a hub located at node k and 0 otherwise, and x_{ijkm} , the fraction of flow sent from node i to node j through hubs k and m in that order. Then the deterministic problem is

(CMAHLP deterministic)

$$\min \sum_{k \in H} f_k y_k + \sum_{(i,j) \in C} \sum_{k \in H} \sum_{m \in H} c_{ijkm} w_{ij} x_{ijkm} \quad (1)$$

$$\text{s.t.} \sum_{k \in H} \sum_{m \in H} x_{ijkm} = 1 \quad \forall (i, j) \in C, \quad (2)$$

$$\sum_{m \in H} x_{ijkm} + \sum_{\substack{m \in H: \\ m \neq k}} x_{ijmk} \leq y_k \quad \forall (i, j) \in C, k \in H, \quad (3)$$

$$\sum_{(i,j) \in C} \sum_{m \in H} w_{ij} x_{ijkm} \leq a_k y_k \quad \forall k \in H, \quad (4)$$

$$y_k \in \{0, 1\} \quad \forall k \in H, \quad (5)$$

$$x_{ijkm} \geq 0 \quad \forall (i, j) \in C, \forall k, m \in H. \quad (6)$$

The objective is to minimize the total cost of opening hubs and transportation costs. Constraints (2) guarantee that pairwise demands are fully satisfied. With constraints (3), direct flow between nonhub nodes is prevented. Constraints (4) are the capacity constraints that limit the total incoming flow at each hub. Constraints (5) and (6) are the domain constraints.

Different from previous studies in the literature, we assume that demand is not known in advance but can be modeled with a polyhedral uncertainty set. We use the hose model introduced by Duffield et al. (1999) and Fingerhut et al. (1997) which is commonly used in the telecommunications literature to represent the demand uncertainty. In this model, instead of estimating pairwise demands, we limit the total flow associated with each demand node. The demand uncertainty set under hose model is

$$D_{\text{hose}} = \left\{ w \in \mathbb{R}_+^{n(n-1)} : \sum_{j \in N \setminus \{i\}} w_{ij} + \sum_{j \in N \setminus \{i\}} w_{ji} \leq b_i, \forall i \in N \right\}, \quad (7)$$

where b_i is the aggregate traffic bound for node $i \in N$. We assume that these bounds are positive and finite for all nodes.

The robust CMAHLP under hose demand uncertainty aims to build a hub network which is viable under any demand realization while minimizing the worst case total cost over all possible demand realizations in the set D_{hose} . Hence the robust problem can be represented as:

$$\begin{aligned} \min & \left(\sum_{k \in H} f_k y_k + \max_{w \in D_{\text{hose}}} \sum_{(i,j) \in C} \sum_{k \in H} \sum_{m \in H} w_{ij} c_{ijkm} x_{ijkm} \right) \\ \text{s.t.} & (2), (3), (5), (6), \\ & \max_{w \in D_{\text{hose}}} \sum_{(i,j) \in C} \sum_{m \in H} w_{ij} x_{ijkm} \leq a_k y_k \quad \forall k \in H. \end{aligned} \quad (8)$$

Here the capacity constraints (4) of the deterministic model are replaced with constraints (8) so that each open hub facility has sufficient capacity to serve under the worst case demand realization in the set D_{hose} .

Observe that this formulation is nonlinear since the demand is a variable. To linearize it, we use a dual transformation, which is widely used in the robust optimization literature (see, e.g., Bertsimas and Sim, 2003 and Altın et al., 2011). For a feasible flow vector x , the inner maximization problem of the objective function,

$$\max_{w \in D_{\text{hose}}} \sum_{(i,j) \in C} \sum_{k \in H} \sum_{m \in H} w_{ij} c_{ijkm} x_{ijkm}, \quad (9)$$

and the maximization problem at the left hand side of the capacity constraint (8),

$$\max_{w \in D_{\text{hose}}} \sum_{(i,j) \in C} \sum_{m \in H} w_{ij} x_{ijkm}, \quad (10)$$

are both linear programming (LP) problems that are feasible and bounded. Therefore the optimal value of these problems are equal to the optimal value of their corresponding duals. Let λ be the dual variable corresponding to the hose model constraint in (7). The dual of problem (9) can be stated as,

$$\min \sum_{i \in N} \lambda_i b_i \quad (11)$$

$$\text{s.t.} \lambda_i + \lambda_j \geq \sum_{k \in H} \sum_{m \in H} c_{ijkm} x_{ijkm} \quad \forall (i, j) \in C, \quad (12)$$

$$\lambda_i \geq 0 \quad \forall i \in N. \quad (13)$$

Similarly, the dual of problem (10) for a given $k \in H$ can be written as follows,

$$\min \sum_{i \in N} \beta_i^k b_i \tag{14}$$

$$\text{s.t. } \beta_i^k + \beta_j^k \geq \sum_{m \in H} x_{ijkm} \quad \forall (i, j) \in C, \forall k \in H, \tag{15}$$

$$\beta_i^k \geq 0 \quad \forall i \in N, \forall k \in H, \tag{16}$$

where β represents the dual variable associated with the hose model constraint. Since these duals are minimization problems, they can be embedded into the original formulation in order to recover linearity. After incorporating these into the robust problem formulation, we obtain the following linear mixed integer programming (MIP) formulation for the robust CMAHLP under hose demand uncertainty:

(CMAHLP hose)

$$\min \sum_{k \in H} f_k y_k + \sum_{i \in N} \lambda_i b_i \tag{17}$$

s.t. (2), (3), (5), (6),

$$\lambda_i + \lambda_j \geq \sum_{k \in H} \sum_{m \in H} c_{ijkm} x_{ijkm} \quad \forall (i, j) \in C, \tag{18}$$

$$\sum_{i \in N} \beta_i^k b_i \leq a_k y_k \quad \forall k \in H, \tag{19}$$

$$\beta_i^k + \beta_j^k \geq \sum_{m \in H} x_{ijkm} \quad \forall (i, j) \in C, \forall k \in H, \tag{20}$$

$$\lambda_i \geq 0 \quad \forall i \in N, \tag{21}$$

$$\beta_i^k \geq 0 \quad \forall i \in N, \forall k \in H. \tag{22}$$

In the deterministic problem, we know that the sum of the capacities of the hubs that are open should be sufficient to satisfy the total demand in the network. In the robust counterpart, we can derive a similar valid inequality by considering the worst case demand.

Theorem 1. Inequality

$$\sum_{k \in H} a_k y_k \geq \min \left\{ \left(\sum_{i \in N} b_i - \max_{i \in N} b_i \right), \sum_{i \in N} b_i / 2 \right\} \tag{23}$$

is a valid inequality.

Proof. It is easy to see that the inequality

$$\sum_{k \in H} a_k y_k \geq \max_{w \in D_{hose}} \sum_{(i,j) \in C} w_{ij}$$

is satisfied by all feasible solutions. The inequality asks to open hubs with sufficient capacity to route the worst case traffic. The right-hand-side of this inequality is an optimization problem. Next we prove that

$$\max_{w \in D_{hose}} \sum_{(i,j) \in C} w_{ij} = \min \left\{ \left(\sum_{i \in N} b_i - \max_{i \in N} b_i \right), \sum_{i \in N} b_i / 2 \right\}.$$

The problem $\max_{w \in D_{hose}} \sum_{(i,j) \in C} w_{ij}$ is

$$\max \sum_{(i,j) \in C} w_{ij}$$

$$\begin{aligned} \text{s.t. } & \sum_{j \in N \setminus \{i\}} w_{ij} + \sum_{j \in N \setminus \{i\}} w_{ji} \leq b_i \quad \forall i \in N, \\ & w_{ij} \geq 0 \quad \forall (i, j) \in C. \end{aligned}$$

Taking the dual of this problem, we obtain the following LP:

$$\begin{aligned} \min & \sum_{i \in N} \vartheta_i b_i \\ \text{s.t. } & \vartheta_i + \vartheta_j \geq 1 \quad \forall (i, j) \in C, \\ & \vartheta_i \geq 0 \quad \forall i \in N. \end{aligned}$$

Observe that the dual problem is the LP relaxation of a weighted vertex covering problem. Nemhauser and Trotter Jr (1974) show that any extreme point ϑ of this LP satisfies $\vartheta_i \in \{0, 1/2, 1\}$ for all $i \in N$. Since we have a covering constraint for all distinct pair of nodes and $b_i > 0$ for all $i \in N$, we can further characterize the optimal solution.

The vector of all ones (1, 1, ..., 1) is clearly not an optimal solution as none of the constraints is tight and one can obtain a better objective function value by decreasing $\vartheta_{i'}$ with $\epsilon > 0$ for an arbitrary $i' \in N$ since $b_{i'}$ is positive. In the case that we know $\vartheta_{i'} = 1/2$ for a node $i' \in N$, $\vartheta_i \geq 1/2$ for all $i \in N \setminus \{i'\}$ for feasibility. Hence the solution with the smallest objective value is the vector (1/2, 1/2, ..., 1/2) with the objective function value equal to $\sum_{i \in N} b_i / 2$. Finally, if there exists a node $i' \in N$ such that $\vartheta_{i'} = 0$, then we must have $\vartheta_i = 1$ for all $i \in N \setminus \{i'\}$ to ensure feasibility. The objective function value of this solution is $\sum_{i \in N} b_i - b_{i'}$. To minimize this value, we set $\vartheta_i = 0$ for a node i with the largest b_i value. Therefore the minimum objective value in this case is $\sum_{i \in N} b_i - \max_{i \in N} b_i$. Hence, the dual optimal value is $\min \{ (\sum_{i \in N} b_i - \max_{i \in N} b_i), \sum_{i \in N} b_i / 2 \}$. By strong duality, this is also the optimal value of the primal. \square

Even though the model CMAHLP hose is a compact linear mixed integer programming model, its size increases rapidly as the number of demand points increases, which makes it difficult to solve for large instances. In the next section, we devise two Benders decomposition algorithms as an attempt to solve large problem instances.

4. Benders reformulations

Benders decomposition is an exact solution method proposed by Benders (1962) and it has been effectively used to solve various mixed integer programming problems in the literature. In this method, the original problem is reformulated by projecting out some of the variables and hence obtaining a formulation with a smaller number of variables and a large number of constraints. One iterates between a master problem, which is a relaxation of the original problem and a subproblem that finds a cut to add to the master problem if the solution of the master problem is not feasible (feasibility cut) or not optimal (optimality cut). In the classical approach, the master is solved to optimality at each iteration. If it is an integer problem, this means that an integer problem is solved from scratch at each iteration. An alternative is to start with a relaxation of small size and solve the reformulation using a cutting plane approach such that each time a candidate solution is found, related cuts are added to the relaxed formulation if the candidate solution is not feasible or optimal. The subproblem is the separation problem solved each time a candidate solution is found. Overall, the problem is solved within one branch-and-cut tree.

The effectiveness of a Benders decomposition algorithm depends on various factors; the number of times the subproblem is solved until optimality is achieved, the computational effort required to solve the master problem and the subproblem etc. In this study, we propose two Benders reformulations for the CMAHLP under hose demand uncertainty by considering these factors in order to obtain an effective decomposition scheme.

4.1. Decomposition by fixing variables y and β (Benders 1)

Consider the mixed integer formulation *CMAHLP hose* as presented in Section 3. Assume that the hub location decisions and the vector β are handled in the master problem and the rest is left to the subproblem. For fixed vectors $y = \hat{y}$ and $\beta = \hat{\beta}$, we obtain the following primal subproblem:

$$(PS1) \quad \min \sum_{i \in N} \lambda_i b_i \quad (24)$$

$$\text{s.t.} \quad \sum_{k \in H} \sum_{m \in H} x_{ijkm} \geq 1 \quad \forall (i, j) \in C, \quad (25)$$

$$\sum_{m \in H} x_{ijkm} + \sum_{\substack{m \in H: \\ m \neq k}} x_{ijmk} \leq \hat{y}_k \quad \forall (i, j) \in C, k \in H, \quad (26)$$

$$\lambda_i + \lambda_j - \sum_{k \in H} \sum_{m \in H} c_{ijkm} x_{ijkm} \geq 0 \quad \forall (i, j) \in C, \quad (27)$$

$$\sum_{m \in H} x_{ijkm} \leq \hat{\beta}_i^k + \hat{\beta}_j^k \quad \forall (i, j) \in C, \forall k \in H, \quad (28)$$

$$\lambda_i \geq 0 \quad \forall i \in N, \quad (29)$$

$$x_{ijkm} \geq 0 \quad \forall (i, j) \in C, \forall k, m \in H. \quad (30)$$

Note that even though we modify constraints (25) here as inequalities, there exists an optimal solution where they hold as equalities. Taking the dual of PS1, we obtain the dual subproblem

$$(DS1) \quad \max \sum_{(i,j) \in C} \rho_{ij} - \sum_{(i,j) \in C} \sum_{k \in H} \hat{y}_k v_{ijk} - \sum_{(i,j) \in C} \sum_{k \in H} (\hat{\beta}_i^k + \hat{\beta}_j^k) u_{ijk} \quad (31)$$

$$\text{s.t.} \quad \sum_{j \in N \setminus \{i\}} w_{ij} + \sum_{j \in N \setminus \{i\}} \omega_{ji} \leq b_i \quad \forall i \in N, \quad (32)$$

$$\rho_{ij} - v_{ijk} - v_{ijm} - u_{ijk} \leq c_{ijkm} \omega_{ij} \quad \forall (i, j) \in C, \forall k, m \in H : k \neq m, \quad (33)$$

$$\rho_{ij} - v_{ijk} - u_{ijk} \leq c_{ijkk} \omega_{ij} \quad \forall (i, j) \in C, \quad \forall k \in H, \quad (34)$$

$$\omega_{ij}, \rho_{ij} \geq 0 \quad \forall (i, j) \in C, \quad (35)$$

$$u_{ijk}, v_{ijk} \geq 0 \quad \forall (i, j) \in C, \forall k \in H, \quad (36)$$

where dual variables ρ, v, ω and u correspond to constraints (25)–(28), respectively. Note that since $\sum_{i \in N} \hat{\beta}_i^k b_i \leq a_k \hat{y}_k$ for all $k \in H$ and $b_i > 0$ for all $i \in N$, we have $\hat{\beta}_i^k = 0$ if $\hat{y}_k = 0$ for all $i \in N$ and $k \in H$. Hence if $\sum_{k \in H} (\hat{\beta}_i^k + \hat{\beta}_j^k) \geq 1$ for all $(i, j) \in C$, the primal subproblem is feasible.

Let S be the set of extreme points (ρ, ω, v, u) of the dual subproblem. Then the master problem can be formulated as follows:

$$(MP1) \quad \min \sum_{k \in H} f_k y_k + q \quad (37)$$

$$\text{s.t.} \quad (5), (19), (22), (23),$$

$$q \geq \sum_{(i,j) \in C} \rho_{ij} - \sum_{(i,j) \in C} \sum_{k \in H} y_k v_{ijk}$$

$$- \sum_{(i,j) \in C} \sum_{k \in H} (\beta_i^k + \beta_j^k) u_{ijk} \quad \forall (\rho, \omega, v, u) \in S, \quad (38)$$

$$\sum_{k \in H} (\beta_i^k + \beta_j^k) \geq 1 \quad \forall (i, j) \in C. \quad (39)$$

Constraints (38) are the Benders optimality cuts and constraints (39) are added to ensure feasibility. In the next subsections we describe how to solve the subproblem efficiently.

4.1.1. Decomposing the subproblem by commodity

In the dual subproblem, constraints (32) and (33)–(34) are interdependent due to the variables ω . In order to eliminate these dependencies, we use the approach by Meraklı and Yaman (2016) and let $\bar{\rho}_{ij} = \frac{\rho_{ij}}{\omega_{ij}}$, $\bar{v}_{ijk} = \frac{v_{ijk}}{\omega_{ij}}$ and $\bar{u}_{ijk} = \frac{u_{ijk}}{\omega_{ij}}$. Then the dual subproblem can be decomposed as

$$\max_{\omega \in D_{hose}} \sum_{(i,j) \in C} w_{ij} \theta_{ij},$$

where for $(i, j) \in C$,

$$(D_{ij}) \quad \theta_{ij} = \max \bar{\rho}_{ij} - \sum_{k \in H} \hat{y}_k \bar{v}_{ijk} - \sum_{k \in H} (\hat{\beta}_i^k + \hat{\beta}_j^k) \bar{u}_{ijk} \quad (40)$$

$$\text{s.t.} \quad \bar{\rho}_{ij} - \bar{v}_{ijk} - \bar{v}_{ijm} - \bar{u}_{ijk} \leq c_{ijkm} \quad \forall k, m \in H : k \neq m, \quad (41)$$

$$\bar{\rho}_{ij} - \bar{v}_{ijk} - \bar{u}_{ijk} \leq c_{ijkk} \quad \forall k \in H, \quad (42)$$

$$\bar{\rho}_{ij} \geq 0, \quad (43)$$

$$\bar{v}_{ijk}, \bar{u}_{ijk} \geq 0 \quad \forall k \in H. \quad (44)$$

The dual of this problem is

$$(P_{ij}) \quad \theta_{ij} = \min \sum_{k \in H} \sum_{m \in H} c_{ijkm} x_{ijkm} \quad (45)$$

$$\text{s.t.} \quad \sum_{k \in H} \sum_{m \in H} x_{ijkm} \geq 1, \quad (46)$$

$$\sum_{m \in H} x_{ijkm} + \sum_{m \in H \setminus \{k\}} x_{ijmk} \leq \hat{y}_k \quad \forall k \in H, \quad (47)$$

$$\sum_{m \in H} x_{ijkm} \leq \hat{\beta}_i^k + \hat{\beta}_j^k \quad \forall k \in H, \quad (48)$$

$$x_{ijkm} \geq 0 \quad \forall k, m \in H. \quad (49)$$

Here it is easy to see that θ_{ij} is the minimum cost of routing commodity $(i, j) \in C$ for given \hat{y} and $\hat{\beta}$ and $\max_{\omega \in D_{hose}} \sum_{(i,j) \in C} w_{ij} \theta_{ij}$ is the worst case cost.

Note that there exists an optimal solution of P_{ij} such that constraint (46) strictly holds. Next we devise an algorithm to compute the optimal dual variables of D_{ij} for any origin destination pair (i, j) .

4.1.2. Computing an optimal solution to problem D_{ij}

For given \hat{y} and $\hat{\beta}$ vectors, the optimal solution of problem P_{ij} can be computed with a simple algorithm. Notice that when all hub capacities are large enough, each flow is routed through the shortest path. In the case of capacitated hubs, this is not necessarily true and the flow sent through a path affects the capacity of the first hub on that path. The flow from i to j using hub k first will

go through only a path $i - k - m_{(k)} - j$ that is a shortest path from i to j using hub k as the first hub, i.e., $m_{(k)} = \operatorname{argmin}_{m \in H: \hat{y}_m=1} c_{ijkm}$ (we pick one arbitrarily in case of multiple minimizers). Besides, $(\hat{\beta}_i^k + \hat{\beta}_j^k)$ value sets a bound on the amount of flow from node i to node j that can be sent through hub k . As the capacity of hub k reserved for commodity (i, j) is known, the routing decision for each commodity becomes independent from each other. Hence, for commodity $(i, j) \in C$, sequencing shortest paths $i - k - m_{(k)} - j$ for each hub k in a nondecreasing order of cost and sending flow from i to j using these paths in a greedy manner provides an optimal solution for our problem.

Algorithm 1 describes how an optimal solution of P_{ij} is com-

Algorithm 1 Compute an optimal solution of P_{ij} .

```

Set  $x_{ijkm} \leftarrow 0 \ \forall k, m \in H$ 
Set  $\text{residual} \leftarrow 1$  and  $p \leftarrow \sum_{k \in H} \hat{y}_k$ 
Sequence hubs as  $k_1, k_2, \dots, k_p$  such that  $c_{ijk_1 m(k_1)} \leq c_{ijk_2 m(k_2)} \leq \dots \leq c_{ijk_p m(k_p)}$ 
for  $h = 1$  to  $p$  do
  if  $\text{residual} > 0$  and  $(\hat{\beta}_i^{k_h} + \hat{\beta}_j^{k_h}) > 0$  then
    Set  $x_{ijk_h m(k_h)} \leftarrow \min\{\text{residual}, (\hat{\beta}_i^{k_h} + \hat{\beta}_j^{k_h})\}$ 
    Set  $\text{residual} \leftarrow \text{residual} - x_{ijk_h m(k_h)}$ 
  end if
end for

```

puted for $(i, j) \in C$. Here *residual* represents the fraction of remaining flow to be sent from node i to node j . Since there exists an optimal solution in which the total fraction of flow sent from i to j is equal to 1, we initially set *residual* to 1. Afterwards, the remaining flow from i to j is routed through hub k with the shortest $i - k - m_{(k)} - j$ path among the hubs that have available capacity.

With the optimal primal solution obtained above, an optimal solution for the dual problem D_{ij} can be constructed using the complementary slackness conditions. An optimal dual solution should satisfy both the constraints (41)-(44) and the complementary slackness conditions given below:

$$\bar{\rho}_{ij} \left(\sum_{k \in H} \sum_{m \in H} x_{ijkm} - 1 \right) = 0 \tag{50}$$

$$\bar{v}_{ijk} \left(\sum_{m \in H} x_{ijkm} + \sum_{m \in H \setminus \{k\}} x_{ijmk} - \hat{y}_k \right) = 0 \quad \forall k \in H, \tag{51}$$

$$\bar{u}_{ijk} \left(\hat{\beta}_i^k + \hat{\beta}_j^k - \sum_{m \in H} x_{ijkm} \right) = 0 \quad \forall k \in H, \tag{52}$$

$$x_{ijkm} (\bar{\rho}_{ij} - \bar{v}_{ijk} - \bar{v}_{ijm} - \bar{u}_{ijk} - c_{ijkm}) = 0 \quad \forall k, m \in H : k \neq m, \tag{53}$$

$$x_{ijkk} (\bar{\rho}_{ij} - \bar{v}_{ijk} - \bar{u}_{ijk} - c_{ijkk}) = 0 \quad \forall k \in H. \tag{54}$$

We compute the dual variables in two steps. First, we fix a set of variables to some feasible values and hence drop the constraints related with them. In the second step, we compute the values of the remaining variables by solving a reduced system of inequalities. At the end, we adjust the variables so that constraints of the dual problem are satisfied. $\kappa \in [0, 1]$ is the scaling parameter used in this adjustment.

The algorithm for computing an optimal solution $(\bar{\rho}_{ij}, \bar{v}_{ij}, \bar{u}_{ij})$ for $(i, j) \in C$ can be seen in Algorithm 2.

Theorem 2. The dual solution computed using Algorithm 2 is optimal for D_{ij} .

Algorithm 2 Compute an optimal solution of D_{ij} .

```

Compute an optimal solution to  $P_{ij}$  using Algorithm 1.
Set  $\bar{\rho}_{ij} = \max_{(k,m) \in F_{ij}} c_{ijkm}$ 
for  $k \in H_1$  do
  Set  $\bar{v}_{ijk} = 0$ 
  if  $\sum_{m \in H} x_{ijkm} > 0$  then
    Set  $\bar{u}_{ijk} = \bar{\rho}_{ij} - c_{ijkm(k)}$ 
  else if  $\sum_{m \in H} x_{ijkm} = 0$  and  $(\hat{\beta}_i^k + \hat{\beta}_j^k) = 0$  then
    Set  $\bar{u}_{ijk} = \max\{0, \bar{\rho}_{ij} - c_{ijkm(k)}\}$ 
  else
    Set  $\bar{u}_{ijk} = 0$ 
  end if
end for
for  $k \in H_0$  do
  Set  $\bar{u}_{ijk} = 0$ 
  Set  $\bar{v}_{ijk} = \max\{0, \max_{m \in H_1} \{\bar{\rho}_{ij} - c_{ijmk} - \bar{u}_{ijm}\}, \max_{m \in H_1} \{\bar{\rho}_{ij} - c_{ijkm}\}, \bar{\rho}_{ij} - c_{ijkk}\}$ 
end for
for  $k, m \in H_0$  such that  $\bar{\rho}_{ij} - c_{ijkm} > 0$  do
  Define  $\Delta = (\bar{\rho}_{ij} - c_{ijkm}) - \bar{v}_{ijk} - \bar{v}_{ijm}$ 
  if  $\Delta > 0$  then
    Update  $\bar{v}_{ijk} \leftarrow \bar{v}_{ijk} + \kappa \Delta$ 
    Update  $\bar{v}_{ijm} \leftarrow \bar{v}_{ijm} + (1 - \kappa) \Delta$ 
  end if
end for

```

Proof. We first check complementary slackness and then dual feasibility.

The dual solution computed using Algorithm 2 satisfies the complementary slackness conditions with the primal solution computed using Algorithm 1. Conditions (50) are satisfied since $\sum_{k \in H} \sum_{m \in H} x_{ijkm} = 1$ for all $(i, j) \in C$. We know that if $\hat{y}_k = 0$, then $\sum_{m \in H} x_{ijkm} + \sum_{m \in H \setminus \{k\}} x_{ijmk} = 0$. If $\hat{y}_k = 1$, i.e., $k \in H_1$, then $\bar{v}_{ijk} = 0$. Hence conditions (51) hold. Conditions (52) are also satisfied. We know that if $\bar{u}_{ijk} > 0$ then $\hat{\beta}_i^k + \hat{\beta}_j^k = 0$ and $\sum_{m \in H} x_{ijkm} = 0$ or path $i - k - m_{(k)} - j$ is used but it is shorter than the longest path among the ones used to send flow from i to j . In the latter case $x_{ijkm(k)} = \hat{\beta}_i^k + \hat{\beta}_j^k$. Hence in both cases the capacity bound on hub k is tight. Therefore if $\bar{u}_{ijk} > 0$ then $\hat{\beta}_i^k + \hat{\beta}_j^k = \sum_{m \in H} x_{ijkm}$. Finally, conditions (53) hold since if $x_{ijkm} > 0$, then $k, m \in H_1$ and thus $\bar{v}_{ijk} = \bar{v}_{ijm} = 0$. In addition, $m = m_{(k)}$ and $\bar{u}_{ijk} = \bar{\rho}_{ij} - c_{ijkm(k)}$. Consequently $\bar{\rho}_{ij} - \bar{v}_{ijk} - \bar{v}_{ijm} - \bar{u}_{ijk} - c_{ijkm} = \bar{\rho}_{ij} - (\bar{\rho}_{ij} - c_{ijkm(k)}) - c_{ijkm(k)} = 0$. We can show that conditions (54) are satisfied in a similar way.

Next we check the dual feasibility of the solution constructed with our algorithm. First we consider the constraints (41). There are four cases:

- **Case 1:** $k \in H_1, m \in H_1$
 Since $k, m \in H_1$ we know that $\bar{v}_{ijk} = \bar{v}_{ijm} = 0$. Hence $\bar{\rho}_{ij} - \bar{v}_{ijk} - \bar{v}_{ijm} - \bar{u}_{ijk} = \bar{\rho}_{ij} - \bar{u}_{ijk}$. We need to consider all possible values of \bar{u}_{ijk} . If $x_{ijkm} > 0$ then $\bar{\rho}_{ij} - \bar{u}_{ijk} = \bar{\rho}_{ij} - (\bar{\rho}_{ij} - c_{ijkm(k)}) = c_{ijkm(k)} = c_{ijkm}$. If $x_{ijkm} = 0$ and $\hat{\beta}_i^k + \hat{\beta}_j^k = 0$, then $\bar{u}_{ijk} = \max\{0, \bar{\rho}_{ij} - c_{ijkm(k)}\}$. Hence $\bar{u}_{ijk} \geq \bar{\rho}_{ij} - c_{ijkm(k)}$ and $\bar{\rho}_{ij} - \bar{u}_{ijk} \leq c_{ijkm(k)}$. We also have $c_{ijkm(k)} \leq c_{ijkm}$ by definition. If $x_{ijkm} = 0$ and $\hat{\beta}_i^k + \hat{\beta}_j^k > 0$, then $\bar{u}_{ijk} = 0$. In this case, we know that $c_{ijkm(k)} \geq \max_{(k',m') \in F_{ij}} c_{ijk'm'}$ since otherwise we would have used the path $i - k - m_{(k)} - j$. As we also have $c_{ijkm} \geq c_{ijkm(k)}$ and $\bar{\rho}_{ij} = \max_{(k',m') \in F_{ij}} c_{ijk'm'}$, we know that (41) is satisfied.
- **Case 2:** $k \in H_1, m \in H_0$

In this case, $\bar{v}_{ijk} = 0$ and $\bar{v}_{ijm} \geq \bar{\rho}_{ij} - c_{ijkm} - \bar{u}_{ijk}$. Then $\bar{\rho}_{ij} - \bar{v}_{ijk} - \bar{v}_{ijm} - \bar{u}_{ijk} = \bar{\rho}_{ij} - \bar{v}_{ijm} - \bar{u}_{ijk} \leq \bar{\rho}_{ij} - \bar{\rho}_{ij} + c_{ijkm} + \bar{u}_{ijk} - \bar{u}_{ijk} = c_{ijkm}$.

• **Case 3:** $k \in H_0, m \in H_1$

In this case, we know that $\bar{v}_{ijm} = \bar{u}_{ijk} = 0$ and $\bar{v}_{ijk} \geq \bar{\rho}_{ij} - c_{ijkm}$. Hence $\bar{\rho}_{ij} - \bar{v}_{ijk} - \bar{v}_{ijm} - \bar{u}_{ijk} \leq \bar{\rho}_{ij} - \bar{\rho}_{ij} + c_{ijkm} \leq c_{ijkm}$.

• **Case 4:** $k \in H_0, m \in H_0$

In this case, we have $\bar{u}_{ijk} = 0$ and $\bar{\rho}_{ij} - c_{ijkm} - \bar{v}_{ijk} - \bar{v}_{ijm} \leq 0$. Hence, $\bar{\rho}_{ij} - \bar{v}_{ijk} - \bar{v}_{ijm} - \bar{u}_{ijk} \leq c_{ijkm}$.

Next we prove that the dual solution satisfies constraints (42). We consider two cases.

• **Case 1:** $k \in H_1$

For $k \in H_1$, the value of \bar{v}_{ijk} is set to zero in our algorithm. Thus, $\bar{\rho}_{ij} - \bar{v}_{ijk} - \bar{u}_{ijk} = \bar{\rho}_{ij} - \bar{u}_{ijk}$. If $x_{ijkk} > 0$ then $\bar{\rho}_{ij} - \bar{u}_{ijk} = \bar{\rho}_{ij} - (\bar{\rho}_{ij} - c_{ijkm(k)}) = c_{ijkk}$ since we are sending flow through path $i - k - j$. When $x_{ijkk} = 0$ and $\hat{\beta}_i^k + \hat{\beta}_j^k = 0$, the value of \bar{u}_{ijk} is set to $\max\{0, \bar{\rho}_{ij} - c_{ijkm(k)}\}$. Then $\bar{\rho}_{ij} - \bar{u}_{ijk} \leq \bar{\rho}_{ij} - (\bar{\rho}_{ij} - c_{ijkm(k)}) = c_{ijkm(k)} \leq c_{ijkk}$. If $x_{ijkk} = 0$ and $\hat{\beta}_i^k + \hat{\beta}_j^k > 0$, then \bar{u}_{ijk} is set to zero. Hence $\bar{\rho}_{ij} - \bar{u}_{ijk} = \bar{\rho}_{ij} = \max_{(k', m') \in F_{ij}} c_{ijk' m'} \leq c_{ijkk}$.

• **Case 2:** $k \in H_0$

In this case, $\bar{u}_{ijk} = 0$ and $\bar{v}_{ijk} \geq \bar{\rho}_{ij} - c_{ijkk}$. Then $\bar{\rho}_{ij} - \bar{v}_{ijk} - \bar{u}_{ijk} \leq \bar{\rho}_{ij} - (\bar{\rho}_{ij} - c_{ijkk}) = c_{ijkk}$.

Since the solution computed using Algorithm 2 is dual feasible and it satisfies complementary slackness conditions with solution x , it is an optimal dual solution. \square

Note that even though we could decompose the dual subproblem into a series of problems, we still generate an aggregate Benders cut.

4.2. Decomposition by projecting out the flow variables (Benders 2)

In this section, we aim to find a decomposition scheme such that the Benders cut can be decomposed for each commodity. For fixed vectors $\hat{y}, \hat{\lambda}$ and $\hat{\beta}$, the subproblem becomes the following feasibility problem:

$$\min 0 \tag{55}$$

s.t. (25), (26), (28), (30),

$$\sum_{k \in H} \sum_{m \in H} c_{ijkm} x_{ijkm} \leq \hat{\lambda}_i + \hat{\lambda}_j \quad \forall (i, j) \in C. \tag{56}$$

For this problem to be feasible, its dual needs to be bounded. So we need

$$\begin{aligned} & \sum_{(i,j) \in C} \rho_{ij} - \sum_{(i,j) \in C} \sum_{k \in H} \hat{y}_k v_{ijk} - \sum_{(i,j) \in C} \sum_{k \in H} (\hat{\beta}_i^k + \hat{\beta}_j^k) u_{ijk} \\ & - \sum_{(i,j) \in C} \omega_{ij} (\hat{\lambda}_i + \hat{\lambda}_j) \leq 0 \end{aligned} \tag{57}$$

for all (ρ, v, ω, u) that satisfy (33) – (36). This system decomposes for each $(i, j) \in C$. Without loss of generality, we can take $\omega_{ij} = 0$ or $\omega_{ij} = 1$ for $(i, j) \in C$. When $\omega_{ij} = 0$, we need

$$\rho_{ij} - \sum_{k \in H} \hat{y}_k v_{ijk} - \sum_{k \in H} (\hat{\beta}_i^k + \hat{\beta}_j^k) u_{ijk} \leq 0$$

for all $(\rho_{ij}, v_{ij}, u_{ij})$ such that

$$\rho_{ij} - v_{ijk} - v_{ijm} - u_{ijk} \leq 0 \quad \forall k, m \in H : k \neq m,$$

$$\rho_{ij} - v_{ijk} - u_{ijk} \leq 0 \quad \forall k \in H,$$

$$\rho_{ij} \geq 0,$$

$$u_{ijk}, v_{ijk} \geq 0 \quad \forall k \in H.$$

It can be seen that this system of inequalities always holds when $\sum_{k \in H} \hat{y}_k \geq 1$ and $\sum_{k \in H} (\hat{\beta}_i^k + \hat{\beta}_j^k) \geq 1$ and the former inequality is already implied by constraint (23). Hence we only need to consider the case $\omega_{ij} = 1$. When we fix $\omega_{ij} = 1$, we obtain

$$\rho_{ij} - \sum_{k \in H} \hat{y}_k v_{ijk} - \sum_{k \in H} (\hat{\beta}_i^k + \hat{\beta}_j^k) u_{ijk} \leq \hat{\lambda}_i + \hat{\lambda}_j$$

for all $(\rho_{ij}, v_{ij}, u_{ij})$ satisfying

$$\rho_{ij} - v_{ijk} - v_{ijm} - u_{ijk} \leq c_{ijkm} \quad \forall (i, j) \in C, \quad \forall k, m \in H : k \neq m, \tag{58}$$

$$\rho_{ij} - v_{ijk} - u_{ijk} \leq c_{ijkm} \quad \forall k \in H, \tag{59}$$

$$\rho_{ij} \geq 0, \tag{60}$$

$$u_{ijk}, v_{ijk} \geq 0 \quad \forall k \in H. \tag{61}$$

Hence, after projecting out x variables, the problem can be reformulated as follows

$$\begin{aligned} & \min \sum_{k \in H} f_k y_k + \sum_{i \in N} \lambda_i b_i \\ & \text{s.t. (5), (19), (21), (22), (23), (39),} \\ & \lambda_i + \lambda_j \geq \rho_{ij} - \sum_{k \in H} y_k v_{ijk} \\ & - \sum_{k \in H} (\beta_i^k + \beta_j^k) u_{ijk} \quad \forall (i, j) \in C, (\rho_{ij}, v_{ij}, u_{ij}) \in S_{ij}, \end{aligned} \tag{62}$$

where S_{ij} is the set of extreme points of the set defined by (58)–(61) for $(i, j) \in C$. The variables corresponding to an extreme point of S_{ij} maximizing the right-hand-side of constraint (62) can be computed as explained in Section 4.1.2.

In this reformulation, we are able to add multiple cuts at each iteration of the Benders decomposition algorithm instead of a single cut since the cuts are disaggregated by commodity.

5. Computational analysis

We test our mathematical model and solution algorithms on well-known Australian Post (AP) and Civil Aeronautics Board (CAB) data set instances with $n = 25, 40, 50$.

The AP data set is first introduced by Ernst and Krishnamoorthy (1996) and it contains postal service data of 200 cities in Australia (accessible from OR-Library, 2015). Each city corresponds to a postal district; city coordinates and pairwise demands are given. The cost multipliers of collection, transfer and distribution are not symmetric; they are taken as $\chi = 3, \alpha = 0.75$ and $\delta = 2$. The pairwise demands are also not symmetric. The demand from a node to itself does not need to be zero. However in our context, we do not allow any demand from a node to itself. To the extend of our knowledge, the AP data set is the only data set with fixed costs and capacities for hubs. For both fixed costs and capacities, two settings are available. Instances with tight (T) fixed costs have larger costs of hub opening compared to the instances with loose (L) fixed costs. Similarly, the instances with tight (T) capacities have smaller available capacities in comparison with the instances with loose (L) capacities. For each problem size n , we consider four cases: LL, LT, TL, TT where the first letter corresponds to the fixed cost setting and the second to the capacity setting.

The CAB data set includes air transportation data for 100 cities in the U.S. (accessible from O’Kelly (1996)). For each city pair, Euclidean distances and demand values are provided. It is assumed

that there is no demand from a node to itself. All distances and demands are symmetric. The cost multipliers of collection and distribution are $\chi = \delta = 1$, respectively. In our experiments, we take the cost multiplier of transfer $\alpha = 0.6, 0.8$. Demand values are scaled so that their sum is equal to one. Unlike the AP data, the CAB data set does not contain information related with fixed costs and capacities. Hence we generated them in the following way: For fixed costs, we considered two different settings L and T where $f_k = 50$ in setting L and $f_k = 100$ in setting T for all possible hub locations $k \in H$. We also randomly generated hub capacities under two settings, L and T , from two different intervals. The hub capacities uniformly take value from interval $[0.5, 0.7]$ in setting L and from interval $[0.4, 0.6]$ in setting T . For all CAB data set instances, we consider four cases LL, LT, TL, TT as for the AP data set.

In our experiments we consider AP and CAB data set instances with $n = 25, 40, 50$. In order to be able to compare our results with benchmark instances, we generated the traffic bounds for the hose model as the sum of nominal demand values associated with each node, i.e., $b_i = \sum_{j \in N \setminus \{i\}} (w_{ij} + w_{ji})$ for all $i \in N$. All nodes are taken as possible hub locations. We perform our computational experiments on a 64-bit machine with Intel Xeon E5-2630 v2 processor at 2.60 GHz and 96 GB of RAM using Java and Cplex 12.5.1. We set a time limit of three hours. All solution times are given in seconds. For the Benders decomposition algorithm implementations, we use the lazy constraint callback function available in CPLEX.

We summarize our computational analysis on the AP and CAB data sets in Tables 1 and 2, respectively. We report the optimal values (the best upper bounds if not solved to optimality), CPU times (the percentage optimality gaps if not solved to optimality) and hub locations for both the deterministic problem and the robust problem. For the robust problem, we compare our results for three different solution methods: the MIP, Benders 1 with single cut approach and Benders 2 with multiple cut approach. We also report the number of cuts added until optimality or time limit is reached for the Benders algorithms. The instances for which we are not able to find an initial solution within three hours of time limit are indicated as *time*. For the instances that we are not able to solve within the time limit but obtain a feasible solution, we report the optimality gaps in brackets. We also mark the instances with more than 100% optimality gaps as *feasible*.

For completeness of analysis, we first present the results of formulation *CMAHLP deterministic* and compare them with the hose model solutions to investigate the effects of demand uncertainty. We use the solutions obtained by Benders 2 since it provides the largest number of optimal solutions to the hose model. Comparing the optimal total costs of the deterministic problem with those of its robust counterpart, we observe a significant increase for both data sets. The optimal values given for the deterministic and robust problems in Table 1 indicate an average increase of 17.09% and a maximum increase of 21.11% in the total costs for the AP data set instances. Similarly, optimal costs of the CAB data set instances given in Table 2 are subject to an increase of 21.02% on average with a maximum increase of 29.71%.

We also compare deterministic and robust problems in terms of optimal hub locations. Considering the AP data set instances which can be solved to optimality for both cases, it can be seen that there is a change in the optimal hub locations in six instances out of nine. In some of them, as in 25LL, only one hub location is changed whereas in some others like 25TL all hubs of the deterministic problem are replaced in the solution of the robust problem. CAB data set instances are more responsive to demand uncertainty. There is a change in the optimal hub locations for all instances except one. For some instances, such as 25LT with $\alpha = 0.6$, these changes are not major. Only the hub at location 23 is moved to location 1. However there are also instances with significant changes in the hub locations. For example, for the instance 25TT

with $\alpha = 0.8$, hub facilities are located at 5, 12 and 21 in the deterministic case, whereas hub locations are at 1 and 22 in the hose model. An interesting observation here is that, while the number of hubs to be opened in the hose model is equal to the number of hubs in the deterministic case for the AP data set instances, it is usually smaller for the CAB data set instances.

Next, we analyze the computational efficiency of our proposed solution methods. In view of our results presented in Tables 1 and 2, it can be seen that even the deterministic problem requires much computational effort using the MIP formulation approach. The AP instances with more than 50 nodes can not be solved to optimality within the time limit. Comparing these results with the ones for *CMAHLP hose*, we observe that the computational effort required to solve the problems to optimality significantly increases with demand uncertainty. The instance 25LL of the AP set can be solved within approximately two hours while it takes five seconds in the deterministic case. For both data sets, only the instances with 25 nodes can be solved to optimality within the time limit.

We also evaluate effectiveness of proposed Benders reformulations. MIP formulation is clearly outperformed by both Benders approaches. All instances that can be solved by the MIP formulation can also be solved using Benders algorithms in less CPU time. The only exception is the CAB data set instance 25LL with $\alpha = 0.6$ for which the MIP formulation performs better than Benders 1, but still Benders 2 has much smaller CPU time. Comparing two Benders decomposition algorithms, we observe that Benders 2 that employs a multiple cut approach is superior to Benders 1 that uses a single cut approach. Our computational results on the AP data set instances can be seen in Table 1. Out of 12 AP data set instances, Benders 1 is able to solve six of them within three hours. Among the instances with 50 nodes, it obtains a feasible solution only for instance 50LT. For the others, it cannot find a feasible solution within the time limit.

On the other hand, Benders 2 succeeds to solve nine instances out of 12 to optimality within the time limit and for the others it is able to obtain feasible solutions with lower optimality gap values than Benders 1. The maximum CPU time for the solved instances is approximately nine minutes. For the CAB data set instances, Benders 2 is able to solve all instances to optimality whereas Benders 1 can not solve six instances out of 24. In five of them, Benders 1 fails to find a feasible solution within the time limit.

Effectiveness and efficiency of Benders algorithms depend mainly on two aspects: computational effort required to solve the subproblems and strength of the optimality and feasibility cuts. Stronger cuts lead to tighter bounds and consequently fewer number of iterations, while a faster solution algorithm for subproblems may significantly reduce the time spent at each iteration. In our implementations, we use Algorithm 2 to find an optimal solution to the dual subproblems. We compare the performance of Algorithm 2 to solving the subproblems with a general purpose solver in Table 3. We solve the AP data instances with Benders 1 and Benders 2 using Algorithm 2 and the corresponding optimization model (D_{ij}), and report the CPU times (the percentage optimality gap if not solved to optimality) and the number of Benders cuts added for each setting. For Benders 1, Algorithm 2 clearly outperforms the optimization model in terms of solution times and the number of instances that can be solved to optimality. This also holds for most instances of Benders 2 except two, 40LT and 40TT. Benders 2 with optimization model is able to solve 40LT to optimality while Algorithm 2 fails, and it takes a smaller CPU time for 40TT. In the light of these results, it can be concluded that Algorithm 2 usually outperforms the optimization model by taking advantage of shorter subproblem solution times.

An important advantage of the hose model is that it requires only the estimation of traffic bounds associated with each node

Table 1
Comparison of solution methods on AP data set instances.

Ins.	CMAHLP deterministic			CMAHLP hose			Hose Benders 1				Hose Benders 2			
	Obj.	cpu(gap)	Hubs	Obj.	cpu(gap)	Hubs	Obj.	cpu(gap)	# Cuts	Hubs	Obj.	cpu(gap)	# Cuts	Hubs
25LL	222411.23	4.26	8,18	269373.49	5896.78	8,19	269373.49	929.93	10,657	8,19	269373.49	5.39	5334	8,19
25LT	248713.51	54.56	9,16,19	299613.46	4883.77	9,12,19	299613.46	2923.96	7002	9,12,19	299613.46	121.20	10,987	9,12,19
25TL	293850.21	26.08	9,23	330504.18	9765.26	11,14	330504.18	40.52	1284	11,14	330504.18	4.33	5058	11,14
25TT	312743.36	227.73	6,14,24	361699.66	3372.66	9,12,14	361699.66	400.15	5180	9,12,14	361699.66	29.78	8806	9,12,14
40LL	230495.10	46.64	14,29	<i>feasible</i>	(100)	<i>H</i>	<i>time</i>	-	-	-	271656.49	53.27	13,446	14,29
40LT	252982.48	909.59	14,26,30	<i>feasible</i>	(100)	<i>H</i>	<i>time</i>	-	-	-	315624.51	(3.22)	63,883	14,26,30
40TL	284821.33	518.83	14,19	<i>feasible</i>	(100)	<i>H</i>	314904.29	285.99	2362	14,19	314904.30	4.74	3852	14,19
40TT	326827.27	2637.96	14,25,38	<i>feasible</i>	(100)	<i>H</i>	<i>time</i>	-	-	-	385661.54	524.30	34,195	14,19,25
50LL	228962.63	176.41	15,35	<i>time</i>	-	-	<i>time</i>	-	-	-	276091.56	285.91	26,848	15,35
50LT	258807.91	(1.04)	6,26,32,48	<i>time</i>	-	-	343860.19	(43.69)	19,164	14,32,35	315039.04	(6.57)	92,034	6,26,32,46
50TL	311199.49	(4.31)	3,45	<i>time</i>	-	-	340552.89	914.53	3362	24,27	340552.90	109.71	21,632	24,27
50TT	395592.49	(6.69)	6,12,26,48	<i>time</i>	-	-	<i>time</i>	-	-	-	452151.20	(11.5)	108,806	25,26,41,48

Table 2
Comparison of solution methods on CAB data set instances.

Ins.	α	CMAHLP deterministic			CMAHLP hose			Hose Benders 1				Hose Benders 2			
		Obj.	cpu(gap)	Hubs	Obj.	cpu(gap)	Hubs	Obj.	cpu(gap)	# Cuts	Hubs	Obj.	cpu	# Cuts	Hubs
25LL	0.6	637.59	4.21	5,14,21,23	776.83	405.54	5,14,21,23	776.83	777.16	10,241	5,14,21,23	776.83	4.08	4343	5,14,21,23
25LT	0.6	637.59	3.38	5,14,21,23	783.75	1630.94	1,5,14,21	783.75	169.91	4295	1,5,14,21	783.75	3.68	4999	1,5,14,21
25TL	0.6	808.92	3.8	4,5,21	937.03	7228.9	22,23	937.03	113.64	1889	22,23	937.03	9.49	6885	22,23
25TT	0.6	808.92	3.84	4,5,21	950.46	7488.49	1,22	950.46	88.00	2134	1,22	950.46	11.85	5844	1,22
25LL	0.8	689.66	3.39	5,12,21	852.38	8072.71	22,23	852.38	327.90	5738	22,23	852.38	12.76	7884	22,23
25LT	0.8	689.66	3.36	5,12,21	860.72	10800.7	1,22	860.72	141.76	3384	1,22	860.72	8.82	5312	1,22
25TL	0.8	831.07	3.5	5,23	952.38	3242.38	22,23	952.38	106.59	2120	22,23	952.38	6.79	5803	22,23
25TT	0.8	839.66	5.11	5,12,21	960.72	2793.97	1,22	960.72	24.93	735	1,22	960.72	4.7	5255	1,22
40LL	0.6	746.22	24.98	5,14,21,23,28,30	2523.06	(100)	<i>H</i>	<i>time</i>	-	-	-	937.44	339.49	21,067	21,22,23,28
40LT	0.6	746.22	26.32	5,14,21,23,28,30	2523.06	(100)	<i>H</i>	<i>time</i>	-	-	-	939.40	338.89	37,823	21,22,23,28
40TL	0.6	926.88	40.69	1,21,28	4523.06	(100)	<i>H</i>	1069.96	1095.26	5137	21,38	1069.96	213.37	30,583	21,38
40TT	0.6	926.88	58.56	1,21,28	4523.06	(100)	<i>H</i>	1079.50	1787.32	8976	21,29	1079.50	183.33	30,962	21,29
40LL	0.8	789.06	24.83	5,21,23,28	2697.41	(100)	<i>H</i>	978.70	1940.47	11,227	21,38	978.70	53.77	13,984	21,38
40LT	0.8	789.06	24.16	5,21,23,28	2697.41	(100)	<i>H</i>	986.81	1409.58	8543	21,29	986.81	59.63	14,878	21,29
40TL	0.8	951.21	786.04	21,23,28	4697.41	(100)	<i>H</i>	1078.70	623.59	4291	21,38	1078.70	41.59	14,783	21,38
40TT	0.8	951.21	723.6	21,23,28	4697.41	(100)	<i>H</i>	1086.81	2375.44	14,269	21,29	1086.81	55.89	15,778	21,29
50LL	0.6	770.46	82.79	5,21,23,28,44	3060.87	(100)	<i>H</i>	<i>time</i>	-	-	-	989.71	1075.26	50,987	21,28,29,44
50LT	0.6	770.46	79.94	5,21,23,28,44	3060.87	(100)	<i>H</i>	<i>time</i>	-	-	-	999.41	3339.56	78,077	21,22,23,28
50TL	0.6	953.53	7977.19	1,21,28	5560.87	(100)	<i>H</i>	1128.81	(7.32)	21,350	29,46	1126.55	1262.27	64,080	21,29
50TT	0.6	953.77	503.1	1,21,28	<i>time</i>	-	-	1147.68	(8.64)	26,198	22,26	1138.50	1869.82	66,096	22,46
50LL	0.8	805.03	72.63	5,21,23,28	<i>time</i>	-	-	<i>time</i>	-	-	-	1036.21	379.4	38,604	21,29
50LT	0.8	805.03	73.16	5,21,23,28	<i>time</i>	-	-	1042.99	8744.93	19,718	22,46	1042.99	367.18	34,822	22,46
50TL	0.8	964.34	3657.09	28,29	<i>time</i>	-	-	1136.21	1618.14	5796	21,29	1136.21	186.25	27,033	21,29
50TT	0.8	977.63	3312.27	1,21,28	<i>time</i>	-	-	1147.23	(4.54)	26,472	21,22	1142.99	118.42	26,332	22,46

Table 3
Comparison of the results of dual algorithm with the optimization model.

Instance	Benders 1				Benders 2			
	Algorithm 2		Optimization Model		Algorithm 2		Optimization Model	
	cpu(gap)	# Cuts	cpu(gap)	# Cuts	cpu(gap)	# Cuts	cpu(gap)	# Cuts
25LL	929.93	10,657	(10.78)	4433	5.39	5334	50.56	4929
25LT	2923.96	7002	(5.73)	4443	121.20	10,987	405.11	19,251
25TL	40.52	1284	3237.61	1339	4.33	5058	50.32	4892
25TT	400.15	5180	4399.50	1860	29.78	8806	105.11	10,421
40LL	time	-	time	-	53.27	13,446	394.87	14,656
40LT	time	-	time	-	(3.22)	63,883	10678.21	66,619
40TL	285.99	2362	time	-	4.74	3852	74.21	3996
40TT	time	-	time	-	524.30	34,195	378.36	13,070
50LL	time	-	time	-	285.91	26,848	1401.7	29,344
50LT	(43.69)	19,164	time	-	(6.57)	92,034	(10.20)	104,887
50TL	914.53	3362	time	-	109.71	21,632	839.38	24,058
50TT	time	-	time	-	(11.50)	108,806	(12.30)	114,049

Table 4
Analysis of changes in the traffic bounds on AP data set instances.

Instance	Deterministic	Hose						
		$\Delta = -0.6$	$\Delta = -0.4$	$\Delta = -0.2$	$\Delta = 0$	$\Delta = 0.2$	$\Delta = 0.4$	$\Delta = 0.6$
25LL	222411.23	125336.88	182522.73	224279.15	269373.49	312357.23	356856.61	393455.53
	8,18	13	12	8,18	8,19	8,19	7,14,17	7,14,17,18
25LT	248713.51	130449.95	192254.46	246613.80	299613.46	352022.55	407825.10	feasible
	9,16,19	12	12,14	9,12,19	9,12,19	2,12,14,19	9,12,14,19	-
25TL	293850.21	141858.38	203274.11	260801.98	330504.18	384527.70	438295.08	487776.48
	9,23	13	14	14	11,14	9,23	6,14,23	6,14,23
25TT	312743.36	145746.23	228548.31	285842.78	361699.66	422022.01	492228.89	564756.80
	6,14,24	14	13,14	12,14	9,12,14	9,12,14	9,12,13,14	6,9,12,14
40LL	230495.10	126956.24	179402.05	227843.65	271656.49	320491.45	363885.20	409324.18
	14,29	19	19	14,29	14,29	10,14,29	6,19,22,29	6,10,21,29
40LT	252982.48	144235.89	199615.74	255202.66	feasible	feasible	feasible	feasible
	14,26,30	14,30	14,30	14,19,30	-	-	-	-
40TL	284821.33	142719.33	195165.15	258651.43	314904.30	376250.91	434153.94	485895.62
	14,19	19	19	14,19	14,19	14,19	19,21	14,19,21
40TT	326827.27	165275.79	227456.35	316004.66	385661.54	471898.40	547132.56	feasible
	14,25,38	14,19	14,19	14,19,35	14,19,25	6,14,19,25	10,14,19,25	-

rather than pairwise demand values. Hence it is worthwhile to consider how the traffic bounds used in the hose model affect the optimal solutions. In Tables 4 and 5, we evaluate the impact of a $100\Delta\%$ change in the estimated traffic bounds for AP and CAB data set instances with $n = 25, 40$, respectively. For each $\Delta \in \{0, \pm 0.2, \pm 0.4, \pm 0.6\}$, we report the optimal total cost and hub locations. The results of $\Delta = 0$ represent our base case where all traffic bounds are correctly estimated. The instances which can not be solved to optimality within the time limit but obtained feasible solutions are excluded from this study and marked as *feasible*. Considering our results, we observed that an increase (decrease) in the traffic bounds usually results in an increase (decrease) in the number of hubs to be opened. However, this is not the case for some such as the CAB data set instances 25LL and 25LT with $\alpha = 0.8$. For these instances the number of hubs to be opened rather increases as we decrease the traffic bounds by 20%. This is mainly due to the fact that the capacity constraints become less tight with a decrease in the traffic bounds and more alternatives for hub selection become available.

Furthermore, it can be seen that the changes in the hub locations are usually in the form of addition and/or removal of hubs. For example, in the optimal solution of CAB data set instance 40LL with $\alpha = 0.6$, hubs are located at 21,22,23,28 for the base case. As traffic bounds are decreased by 20%, the hub at 21 and 28 are removed whereas after an increase of 20% additional hubs are located at nodes 14 and 30.

Another aspect to be evaluated is how the deterministic problem solutions perform under worst case demand scenario. We

present our findings in Tables 6 and 7 for the AP and CAB data sets with $n = 25, 40$. For each traffic bound level $\Delta \in \{0, 0.2, 0.4, 0.6\}$, we reported the total costs for the case in which hubs are located as in the deterministic problem solutions and the percentage deviation from the optimal value of the hose model with $100\Delta\%$ increase in the traffic bounds. We do not consider negative Δ values since our main concern is the feasibility of the deterministic solutions and when $\Delta \leq 0$, deterministic solutions are always feasible. The instances for which deterministic solutions are not feasible are indicated as *infeasible*. For the instances that we are not able to obtain the optimal solution to the hose model, we report the objective value of the deterministic solution but mark the deviation with -. We observe that AP data set instances are more sensitive to the changes in the traffic bounds. When traffic bounds are increased by 20% for the AP data set, the solutions of the deterministic problem become infeasible for four out of eight instances. On the other hand, even after a 60% increase in the traffic bounds, the deterministic solution is not feasible in only one instance of the CAB data set. The reason behind this may be that the capacity constraints in the AP data set are much tighter than in the CAB data set. These results show that using deterministic formulations for the CMAHLP may cause infeasibilities because of the demand uncertainty. Moreover, ignoring uncertainty in the demand may also result in a significant increase in the total costs. By incorporating demand uncertainty into our decisions, we may save up to 8.84% for the AP data set and 14.50% for the CAB data set instances.

Table 5
Analysis of changes in the traffic bounds on CAB data set instances.

Instance	α	Deterministic	Hose						
			$\Delta = -0.6$	$\Delta = -0.4$	$\Delta = -0.2$	$\Delta = 0$	$\Delta = 0.2$	$\Delta = 0.4$	$\Delta = 0.6$
25LL	0.6	637.59	357.63	514.57	661.25	776.83	913.16	1050.32	1180.01
		5,14,21,23	1	23	5,14,21,23	5,14,21,23	5,14,21,23	5,14,21,22,23	5,14,21,22,23
25LT	0.6	637.59	357.63	523.78	658.64	783.75	918.30	1052.39	1183.21
		5,14,21,23	1	1,21	1,5,14,21	1,5,14,21	1,5,14,21	1,5,14,21,22	1,5,9,14,21,22
25TL	0.6	808.92	407.63	564.57	777.86	937.03	1102.39	1266.80	1417.51
		4,5,21	1	23	22,23	22,23	22,23	14,20,21,23	14,21,22,23
25TT	0.6	808.92	407.63	623.78	788.02	950.46	1118.30	1275.30	1427.42
		4,5,21	1	1,21	1,21	1,22	1,5,14,21	1,2,5,21	1,14,18,21
25LL	0.8	689.66	357.63	514.57	782.71	852.38	1015.59	1165.83	1314.90
		5,12,21	1	23	23,28,38	22,23	22,23	14,21,22,23	14,21,22,23
25LT	0.8	689.66	357.63	542.69	786.47	860.72	1019.50	1167.22	1317.79
		5,12,21	1	1,21	21,28,29	1,22	1,18,21	1,14,21,22	1,14,18,21
25TL	0.8	831.07	407.63	564.57	888.16	952.38	1115.59	1319.36	1473.13
		5,23	1	23	21,29	22,23	22,23	18,21,23	18,21,23
25TT	0.8	839.66	407.63	642.69	890.91	960.72	1169.50	1322.06	1482.03
		5,12,21	1	1,21	21,29	1,22	1,18,21	1,18,21	1,18,21
40LL	0.6	746.22	405.33	583.00	695.16	937.44	1078.75	1217.40	1355.74
		5,14,21,23,28,30	38	38	22,23	21,22,23,28	14,21,22,23,28,30	5,14,21,23,28,29,30	5,14,21,23,28,30,38
40LT	0.6	746.22	405.33	609.47	700.69	939.40	1082.32	1225.66	1374.00
		5,14,21,23,28,30	38	21,29	1,21	21,22,23,28	14,21,22,23,28,30	5,14,21,23,28,29,30	5,14,21,28,29,30,32
40TL	0.6	926.88	455.33	633.00	795.16	1069.96	1252.82	1452.52	1628.58
		1,21,28	38	38	22,23	21,38	21,38	22,28,38	21,22,23,28
40TT	0.6	926.88	455.33	709.47	800.69	1079.50	1286.48	<i>feasible</i>	1637.10
		1,21,28	38	21,29	1,21	21,29	21,22,23	-	21,22,23,28
40LL	0.8	789.06	405.33	583.00	796.73	978.70	1162.42	1345.95	<i>feasible</i>
		5,21,23,28	38	38	21,29	21,38	21,38	21,22,23	-
40LT	0.8	789.06	405.33	620.17	798.13	986.81	1173.21	1351.77	1532.80
		5,21,23,28	38	21,29	21,29	21,29	21,22,23	21,22,23	21,22,23
40TL	0.8	951.21	455.33	633.00	896.73	1078.70	1262.42	1495.95	1673.83
		21,23,28	38	38	21,29	21,38	21,38	21,22,23	21,22,23
40TT	0.8	951.21	455.33	720.17	898.13	1086.81	1323.21	1501.77	1682.80
		21,23,28	38	21,29	21,29	21,29	21,22,23	21,22,23	21,22,23

Table 6
Cost and feasibility analysis on AP data set instances.

Instance	Deterministic	Hose			
		$\Delta = 0$	$\Delta = 0.2$	$\Delta = 0.4$	$\Delta = 0.6$
25LL	8,18	273689.90	328821.59	388400.77	<i>Infeasible</i>
		1.60	5.27	8.84	-
25LT	9,16,19	318566.37	<i>Infeasible</i>	<i>Infeasible</i>	<i>Infeasible</i>
		6.33	-	-	-
25TL	9,23	338344.95	384527.70	439179.37	498074.00
		2.37	0.00	0.20	2.11
25TT	6,14,24	380116.35	<i>Infeasible</i>	<i>Infeasible</i>	<i>Infeasible</i>
		5.09	-	-	-
40LL	14,29	271656.49	321745.66	381296.09	<i>Infeasible</i>
		0.00	0.39	4.78	-
40LT	14,26,30	315624.51	<i>Infeasible</i>	<i>Infeasible</i>	<i>Infeasible</i>
		-	-	-	-
40TL	14,19	314904.30	376250.91	441187.27	<i>Infeasible</i>
		0.00	0.00	1.62	-
40TT	14,25,38	409077.25	<i>Infeasible</i>	<i>Infeasible</i>	<i>Infeasible</i>
		6.07	-	-	-

6. Conclusions

In this study, we considered a capacitated multiple allocation hub location problem where the demand can take any value from a hose uncertainty set. We proposed a mixed integer programming formulation and devised two different Benders decomposition algorithms. In view of our computational results, Benders 2, which utilizes a multicut approach outperformed the other approaches in terms of computational efficiency. It succeeded to solve AP data set instances with up to 50 nodes and obtained feasible solutions with relatively lower optimality gaps for the instances that could

not be solved within the time limit. Our computational results also showed that ignoring uncertainty may lead to solutions in which congestion can occur at hubs as a result of exceeded capacities and routing costs can be significantly higher for some demand realizations. We also observed that, unlike the uncapacitated problem, there may be significant changes in hub locations when demand uncertainty and hub capacities are considered simultaneously.

The multiple allocation hub location problem is based on many underlying assumptions such as a complete hub network, zero fixed costs for connections and the same discount factor for all hub to hub connections. Relaxing one or more of these assump-

Table 7
Cost and feasibility analysis on CAB data set instances.

Instance	α	Deterministic	Hose			
			$\Delta = 0$	$\Delta = 0.2$	$\Delta = 0.4$	$\Delta = 0.6$
25LL	0.6	5,14,21,23	776.83	913.16	1068.37	1231.41
			0.00	0.00	1.72	4.36
25LT	0.6	5,14,21,23	804.26	944.93	1105.18	1269.00
			2.62	2.90	5.02	7.25
25TL	0.6	4,5,21	1000.66	1178.92	1364.74	1563.07
			6.79	6.94	7.73	10.27
25TT	0.6	4,5,21	988.62	1163.36	1343.69	1533.70
			4.02	4.03	5.36	7.45
25LL	0.8	5,12,21	925.62	1101.89	1284.74	1500.92
			8.59	8.50	10.20	14.15
25LT	0.8	5,12,21	917.15	1091.15	1267.58	1459.95
			6.56	7.03	8.60	10.79
25TL	0.8	5,23	984.60	1203.44	<i>Infeasible</i>	<i>Infeasible</i>
			3.38	7.87	-	-
25TT	0.8	5,12,21	1067.15	1241.15	1417.58	1609.95
			11.08	6.13	7.23	8.63
40LL	0.6	5,14,21,23,28,30	954.45	1108.71	1275.14	1453.83
			1.81	2.78	4.74	7.24
40LT	0.6	5,14,21,23,28,30	957.85	1121.26	1293.44	1481.57
			1.96	3.60	5.53	7.83
40TL	0.6	1,21,28	1119.16	1301.74	1526.82	1794.77
			4.60	3.90	5.11	10.20
40TT	0.6	1,21,28	1122.54	1324.41	1570.80	1874.42
			3.99	2.95	-	14.50
40LL	0.8	5,21,23,28	1041.52	1219.70	1406.76	1601.52
			6.42	4.93	4.52	-
40LT	0.8	5,21,23,28	1042.70	1224.21	1415.51	1624.80
			5.66	4.35	4.72	6.00
40TL	0.8	21,23,28	1174.57	1361.55	1570.35	1840.42
			8.89	7.85	4.97	9.95
40TT	0.8	21,23,28	1176.62	1371.64	1614.60	1932.29
			8.26	3.66	7.51	14.83

tions results in a more realistic and challenging problem. As future research, we are interested in incorporating demand uncertainty into these problems.

Acknowledgment

The research of the first author is supported by the Scientific and Technological Research Council of Turkey (TUBITAK), Grant Number: BIDEB-2211-A, and the research of the second author is supported by the Turkish Academy of Sciences.

References

Altın, A., Amaldi, E., Belotti, P., Pınar, M., 2007. Provisioning virtual private networks under traffic uncertainty. *Networks* 49 (1), 100–115.

Altın, A., Yaman, H., Pınar, M.Ç., 2011. The robust network loading problem under hose demand uncertainty: formulation, polyhedral analysis, and computations. *INFORMS J. Comput.* 23 (1), 75–89.

Alumur, S., Kara, B.Y., 2008. Network hub location problems: the state of the art. *Eur. J. Oper. Res.* 190 (1), 1–21.

Alumur, S.A., Nickel, S., da Gama, F.S., 2012. Hub location under uncertainty. *Transp. Res. B* 46 (4), 529–543.

Atamtürk, A., 2006. Strong formulations of robust mixed 0–1 programming. *Math. Program.* 108 (2–3), 235–250.

Ben-Tal, A., Goryashko, A., Guslitzer, E., Nemirovski, A., 2004. Adjustable robust solutions of uncertain linear programs. *Math. Program.* 99 (2), 351–376.

Ben-Tal, A., Nemirovski, A., 1998. Robust convex optimization. *Math. Oper. Res.* 23 (4), 769–805.

Ben-Tal, A., Nemirovski, A., 1999. Robust solutions of uncertain linear programs. *Oper. Res. Lett.* 25 (1), 1–13.

Ben-Tal, A., Nemirovski, A., 2008. Selected topics in robust convex optimization. *Math. Program.* 112 (1), 125–158.

Benders, J.F., 1962. Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik* 4 (1), 238–252.

Bertsimas, D., Sim, M., 2003. Robust discrete optimization and network flows. *Math. Program.* 98 (1–3), 49–71.

Bertsimas, D., Sim, M., 2004. The price of robustness. *Oper. Res.* 52 (1), 35–53.

Boland, N., Krishnamoorthy, M., Ernst, A.T., Ebery, J., 2004. Preprocessing and cutting for multiple allocation hub location problems. *Eur. J. Oper. Res.* 155 (3), 638–653.

Camargo, R., d. Mirandajr., G., Luna, H.P., 2008. Benders decomposition for the uncapacitated multiple allocation hub location problem. *Comput. Oper. Res.* 35 (4), 1047–1064.

Camargo, R., d. Mirandajr., G., Luna, H.P., 2009. Benders decomposition for hub location problems with economies of scale. *Transp. Sci.* 43 (1), 86–97.

Campbell, J.F., 1992. Location and allocation for distribution systems with transshipments and transportation economies of scale. *Ann. Oper. Res.* 40 (1), 77–99.

Campbell, J.F., 1994. Integer programming formulations of discrete hub location problems. *Eur. J. Oper. Res.* 72 (2), 387–405.

Campbell, J.F., 1994. A survey of network hub location. *Stud. Locat. Anal.* 6, 31–49.

Campbell, J.F., 1996. Hub location and the p -hub median problem. *Oper. Res.* 44 (6), 923–935.

Campbell, J.F., Ernst, A.T., Krishnamoorthy, M., 2002. Hub location problems. *Facil. Locat.: Appl. Theory* 1, 373–407.

Campbell, J.F., O’Kelly, M.E., 2012. Twenty-five years of hub location research. *Transp. Sci.* 46 (2), 153–169.

Cánovas, L., García, S., Marín, A., 2007. Solving the uncapacitated multiple allocation hub location problem by means of a dual-ascent technique. *Eur. J. Oper. Res.* 179 (3), 990–1007.

Chekuri, C., Shepherd, F., Oriolo, G., Scutellà, M., 2007. Hardness of robust network design. *Networks* 50 (1), 50–54.

Contreras, I., Cordeau, J.-F., Laporte, G., 2011. Benders decomposition for large-scale uncapacitated hub location. *Oper. Res.* 59 (6), 1477–1490.

Contreras, I., Cordeau, J.-F., Laporte, G., 2011. Stochastic uncapacitated hub location. *Eur. J. Oper. Res.* 212 (3), 518–528.

Contreras, I., Cordeau, J.-F., Laporte, G., 2012. Exact solution of large-scale hub location problems with multiple capacity levels. *Transp. Sci.* 46 (4), 439–459.

Duffield, N.G., Goyal, P., Greenberg, A., Mishra, P., Ramakrishnan, K.K., van der Merive, J.E., 1999. A flexible model for resource management in virtual private networks. In: *ACM SIGCOMM Computer Communication Rev.*, 29. ACM, pp. 95–108.

Ebery, J., Krishnamoorthy, M., Ernst, A.T., Boland, N., 2000. The capacitated multiple allocation hub location problem: formulations and algorithms. *Eur. J. Oper. Res.* 120 (3), 614–631.

Ernst, A.T., Krishnamoorthy, M., 1996. Efficient algorithms for the uncapacitated single allocation p -hub median problem. *Locat. Sci.* 4 (3), 139–154.

Ernst, A.T., Krishnamoorthy, M., 1998a. Exact and heuristic algorithms for the uncapacitated multiple allocation p -hub median problem. *Eur. J. Oper. Res.* 104 (1), 100–112.

Ernst, A.T., Krishnamoorthy, M., 1998b. An exact solution approach based on shortest-paths for p -hub median problems. *INFORMS J. Comput.* 10 (2), 149–162.

Farahani, R.Z., Hekmatfar, M., Arabani, A.B., Nikbaksh, E., 2013. Hub location problems: a review of models, classification, solution techniques, and applications. *Comput. Ind. Eng.* 64 (4), 1096–1109.

- Fingerhut, J.A., Suri, S., Turner, J.S., 1997. Designing least-cost nonblocking broadband networks. *J. Algorithms* 24 (2), 287–309.
- Gelareh, S., Nickel, S., 2011. Hub location problems in transportation networks. *Transp. Res. E* 47 (6), 1092–1111.
- Hamacher, H.W., Labbé, M., Nickel, S., Sonneborn, T., 2004. Adapting polyhedral properties from facility to hub location problems. *Discrete Appl. Math.* 145 (1), 104–116.
- Italiano, G., Leonardi, S., Oriolo, G., 2006. Design of trees in the hose model: the balanced case. *Oper. Res. Lett.* 34 (6), 601–606.
- Klincewicz, J.G., 1996. A dual algorithm for the uncapacitated hub location problem. *Locat. Sci.* 4 (3), 173–184.
- Klincewicz, J.G., 1998. Hub location in backbone/tributary network design: a review. *Locat. Sci.* 6 (1), 307–335.
- Marianov, V., Serra, D., 2003. Location models for airline hubs behaving as m/d/c queues. *Comput. Oper. Res.* 30 (7), 983–1003.
- Marín, A., 2005a. Formulating and solving splittable capacitated multiple allocation hub location problems. *Comput. Oper. Res.* 32 (12), 3093–3109.
- Marín, A., 2005b. Uncapacitated Euclidean hub location: strengthened formulation, new facets and a relax-and-cut algorithm. *J. Glob. Optim.* 33 (3), 393–422.
- Mayer, G., Wagner, B., 2002. Hublocator: an exact solution method for the multiple allocation hub location problem. *Comput. Oper. Res.* 29 (6), 715–739.
- Meraklı, M., Yaman, H., 2016. Robust intermodal hub location under polyhedral demand uncertainty. *Transp. Res. B* 86, 66–85.
- Mudchanatongsuk, S., Ordóñez, F., Liu, J., 2008. Robust solutions for network design under transportation cost and demand uncertainty. *J. Oper. Res. Soc.* 59 (5), 652–662.
- Nemhauser, G.L., Trotter Jr, L.E., 1974. Properties of vertex packing and independence system polyhedra. *Math. Program.* 6 (1), 48–61.
- O'Kelly, M. E., 1996. Cab100. [Online; accessed 28-August-2015].
- O'Kelly, M.E., Miller, H.J., 1994. The hub network design problem: a review and synthesis. *J. Transp. Geogr.* 2 (1), 31–40.
- OR-Library, 2015.
- Ordóñez, F., Zhao, J., 2007. Robust capacity expansion of network flows. *Networks* 50 (2), 136–145.
- Rodríguez-Martín, I., Salazar-González, J.J., 2008. Solving a capacitated hub location problem. *Eur. J. Oper. Res.* 184 (2), 468–479.
- Sasaki, M., Fukushima, M., 2003. On the hub-and-spoke model with arc capacity constraints. *J. Oper. Res. Soc. of Japan-Keiei Kagaku* 46 (4), 409–428.
- Shahabi, M., Unnikrishnan, A., 2014. Robust hub network design problem. *Transp. Res. E* 70, 356–373.
- Sim, T., Lowe, T.J., Thomas, B.W., 2009. The stochastic p -hub center problem with service-level constraints. *Comput. Oper. Res.* 36 (12), 3166–3177.
- Skorin-Kapov, D., Skorin-Kapov, J., O'Kelly, M., 1996. Tight linear programming relaxations of uncapacitated p -hub median problems. *Eur. J. Oper. Res.* 94 (3), 582–593.
- Yaman, H., Karaşan, O.E., Pınar, M.Ç., 2001. The robust spanning tree problem with interval data. *Oper. Res. Lett.* 29 (1), 31–40.
- Yaman, H., Karaşan, O.E., Pınar, M.Ç., 2007. Restricted robust uniform matroid maximization under interval uncertainty. *Math. Program.* 110 (2), 431–441.
- Yang, T.-H., 2009. Stochastic air freight hub location and flight routes planning. *Appl. Math. Modell.* 33 (12), 4424–4430.