



Benders Decomposition Algorithms for Two Variants of the Single Allocation Hub Location Problem

Nader Ghaffarinasab¹ · Bahar Y. Kara²

Published online: 9 October 2018

© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract

The hub location problem (HLP) is a special type of the facility location problem with numerous applications in the airline industry, postal services, and computer and telecommunications networks. This paper addresses two basic variants of the HLP, namely the uncapacitated single allocation hub location problem (USAHLP) and the uncapacitated single allocation p -hub median problem (USAPHMP). Exact solution procedures based on Benders decomposition algorithm are proposed to tackle large sized instances of these problems. The standard Benders decomposition algorithm is enhanced through implementation of several algorithmic refinements such as using a new cut disaggregation scheme, generating strong optimality cuts, and an efficient algorithm to solve the dual subproblems. Furthermore, a modern implementation of the algorithm is used where a single search tree is established for the problem and Benders cuts are successively added within a branch-and-cut framework. Extensive computational experiments are conducted to examine the efficiency of the proposed algorithms. We have been able to solve all the instances of the problems from all three main data sets of the HLP to optimality in reasonable computational times.

Keywords Hub location problem · Single allocation · Benders decomposition · Branch-and-cut · Algorithmic refinements

1 Introduction

Hubs are intermediate facilities in many-to-many distribution systems where key operations such as transshipment, consolidation, break-bulk, etc. are performed. Despite

✉ Nader Ghaffarinasab
ngnasab@tabrizu.ac.ir

Bahar Y. Kara
bkara@bilkent.edu.tr

¹ Department of Industrial Engineering, University of Tabriz, Tabriz, Iran

² Department of Industrial Engineering, Bilkent University, Ankara, Turkey

the traditional point-to-point shipment paradigm, traffic associated with origin–destination (O/D) pairs in the hub networks are routed via hub facilities resulting in smaller number of linkages each carrying a relatively large flow volume. The resulting concentration of traffic flows makes it possible to take advantage of economies of scale by efficiently utilizing the transportation capacities especially on inter-hub connections. The hub location problem (HLP) consists of determining the locations of hub facilities and allocating the demand nodes to these hubs in such a manner that an objective function of interest (usually total system-wide costs) is optimized.

Having determined the location of hubs in the network, the nodes which are not selected as hubs can be assigned to the installed hubs according to different allocation protocols. Two of the main allocation types are the single and multiple allocation schemes. Based on the single allocation protocol, which is the underlying network topology in the current work, all the incoming/outgoing traffic to/from a non-hub node is routed through a single hub, whereas in the multiple allocation scheme, each non-hub node can receive and send flow through more than one hub. Figure 1 illustrates an example of single allocation hub network in which the square and circular nodes represent the hubs and the non-hub nodes, respectively. From an application point of view, single allocation networks are needed in several contexts such as the less-than-truckload (LTL) trucking industry where each end-of-line terminal might be assigned to a single break-bulk terminal. In a similar manner, in the telecommunication industry, when the objective is to reduce the cost of constructing the network, single allocation networks are usually preferred (Campbell and O’Kelly 2012).

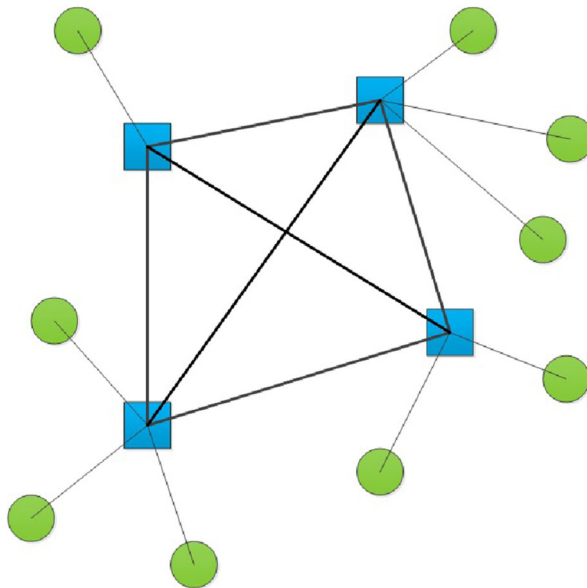


Fig. 1 An example of single allocation hub network

HLPs constitute a difficult class of NP-hard combinatorial optimization problems (Contreras et al. 2011a). Moreover, in case of the single allocation HLP, given a fixed set of locations for the hubs, the allocation part of the problem is still NP-hard (Kara 1999) making it very difficult, if not impossible, to solve the real-world (usually large-sized) instances of the problem to optimality using standard optimization packages. For this reason, developing efficient solution algorithms capable of solving the problem instances of large sizes is of utmost practical importance.

In this paper, we address two of the basic variants of the single allocation HLP, namely the uncapacitated single allocation hub location problem (USAHLP) and the uncapacitated single allocation p -hub median problem (USA p HMP) and propose exact solution algorithms for solving these problems. The USAHLP assumes that the number of hubs to be opened in the network is endogenously determined by solving the mathematical model in which the objective is to minimize the total fixed hub establishment costs as well as variable transportation costs. In contrast, the number of hubs to be opened in the USA p HMP is an input parameter to the model and the objective is to minimize only the transportation costs.

The main contribution of this paper is to propose exact algorithms based on Benders decomposition for solving large-scale instances of the USAHLP and USA p HMP. The standard implementation of the algorithm is enhanced by using a new cut disaggregation scheme, generating strong optimality cuts, and an efficient algorithm to solve the dual subproblems. In order to evaluate the efficiency of the proposed algorithms, an extensive set of computational experiments were conducted on all the three well-known data sets available in the literature of hub location. It was shown that, the proposed algorithms are able to solve large-scale instances of the problems to optimality in quite short and reasonable computational times.

The remainder of this paper is organized as follows. The next section discusses the relevant literature for the HLP, the USAHLP, and the USA p HMP. In Section 3, we present the MIP formulations for the USAHLP and USA p HMP. The proposed Benders decomposition algorithms along with the associated algorithmic refinements are presented in Section 4. Computational experiments as well as the corresponding results are presented in Section 5. Finally, Section 6 provides some conclusions and outlooks for future research.

2 Literature Review

The USAHLP was first introduced by O'Kelly (1992) and formulated as a quadratic integer program. Over the years, several studies have considered the USAHLP. Campbell (1994) presented the first linear programming formulations for the uncapacitated single and multiple allocation HLPs. Abdinnour-Helm and Venkataramanan (1998) presented a branch-and-bound (B&B) procedure as well as a genetic algorithm (GA) to solve the problem. Their (B&B) could solve problem instances of less than 20 nodes, whereas they were able to solve instances with up to 80 nodes using the proposed GA. Topcuoglu et al. (2005) proposed a GA for the USAHLP and showed that their algorithm significantly surpassed the related work prior to that time.

Another heuristic for this problem was proposed by Chen (2007). The author proposed two approaches to determine upper bounds for the number of hubs along with a hybrid heuristic based on the simulated annealing (SA), tabu list, and improvement procedures. It was shown that the proposed heuristic outperformed the algorithms presented in Topcuoglu et al. (2005) in terms of running time and solution quality.

Silva and Cunha (2009) proposed three multi-start tabu search (TS) algorithm and a two-stage integrated TS heuristic for the USAHLP and solved instances with up to 200 nodes using their heuristics. Marić et al. (2013) presented a memetic algorithm (MA) for solving the USAHLP. Their algorithm used two efficient local search heuristics in order to improve both the location and allocation part of the problem. A new efficient TS algorithm for the USAHLP was proposed by Abyazi-Sani and Ghanbari (2016). The authors compared their TS algorithm with that of Silva and Cunha (2009) and showed that their algorithm was able to find best known solutions for all standard test problems in short computational time. Meier et al. (2016) and Meier and Clausen (2017) proposed a row generation (RG) algorithm for solving the single allocation HLPs assuming that the transportation costs between the hubs are proportional to the corresponding Euclidean distances. This assumption enabled them to construct a new linearization method for the quadratic terms of the single allocation HLP models and incorporate it into the RG procedure for solving instances of up to 200 nodes to optimality.

Some authors have addressed the capacitated version of the single allocation hub location problem where the total ingoing/outgoing flow to/from each installed hub is not allowed to exceed a pre-specified threshold level. Aykin (1994) considered the capacitated single allocation hub location problem (CSAHLP) in which the direct service was allowed between non-hub nodes. A branch-and-bound algorithm and a heuristic procedure partitioning the set of solutions on the basis of hub locations were presented. Problem instances of up to 40 nodes were solved by the proposed algorithms. Ernst and Krishnamoorthy (1999) presented an SA and a random descent algorithm for solving the CSAHLP. Using the proposed algorithms they found near-optimal solutions for the problems of up to 200 nodes. Contreras et al. (2009) proposed a Lagrangian relaxation algorithm to obtain tight upper and lower bounds for the CSAHLP. Contreras et al. (2011c) presented a branch-and-price algorithm for solving the CSAHLP where they used the Lagrangian relaxation developed in Contreras et al. (2009) to obtain tight lower bounds of the restricted master problem. Instances of up to 200 nodes were solved to optimality using the proposed algorithm.

The USA p HMP was first formulated by O'Kelly (1987) as a quadratic integer program. He proposed two heuristic solution methods to solve the problem. Later, Klincewicz (1992) developed two heuristics based on TS and greedy randomized search procedure (GRASP) for the USA p HMP and solved instances with up to 52 nodes using the proposed heuristics. Ernst and Krishnamoorthy (1996) proposed an SA algorithm that used two move strategies for generating new solutions. They provided solutions for the instances with up to 200 nodes. Skorin-Kapov and Skorin-Kapov (1994) developed a TS algorithm for the USA p HMP and solved instances of 25 nodes with their heuristic. Skorin-Kapov et al. (1996) developed new mixed linear formulations with tight linear programming relaxations for the USA p HMP and

its multiple allocation version. Also, for the single allocation case, they were able to establish optimality of all heuristic solutions obtained via the tabu search algorithm from their previous study (Skorin-Kapov and Skorin-Kapov 1994).

Kratika et al. (2007) developed two GA-based approaches for the problem that used a set of mechanisms to increase the diversity of solutions. They improved the best-known solutions from literature for large scale instances with up to 200 nodes. Ernst and Krishnamoorthy (1998) proposed a branch-and-bound algorithm which started with a set of root nodes rather than a single root node. Computational experiments revealed that their algorithm was very efficient for small values of p . The authors solved test instances of up to 100 nodes and with $p = 2$ and $p = 3$. A hybrid algorithm combining evolutionary algorithms and local search was developed by Wolf and Merz (2007). More recently, Ilić et al. (2010) presented a new general variable neighborhood search (GVNS) approach for the USA p HMP. The experimental results showed that the GVNS based heuristics outperformed the best-known heuristics in terms of solution quality and computational effort. Peker et al. (2016) identified general features of optimal hub locations for the USA p HMP and then exploited this knowledge to develop a straightforward heuristic methodology for solving the problem. Taner and Kara (2016) proposed a new approach to forecast the change in the demand patterns generated at different locations as a result of the placement of new hubs in the USA p HMP. The interested readers are referred to the papers (Alumur and Kara 2008; Campbell and O’Kelly 2012; Contreras 2015; Farahani et al. 2013) as recent and detailed surveys on the HLP.

Benders decomposition (Benders 1962) is a row generation based method for solving mixed integer programming (MIP) problems. The approach is based on a relaxation algorithm for solving the problem by partitioning it into two simpler problems. For the facility location and network design problems, Benders decomposition algorithm has been successfully applied in numerous research works such as Fazel-Zarandi et al. (2013), Geoffrion and Graves (1974), Magnanti and Wong (1981), Tang et al. (2013), Tang et al. (2016), Wentges (1996), and He et al. (2018), among others. In case of HLP, many researchers have successfully designed and applied Benders decomposition algorithms to solve the problem in multiple allocation setting (see e.g., de Camargo et al. 2008, 2009b, b; Contreras et al. 2011a, b, 2012; Gelareh and Nickel 2008, 2011; Merakli and Yaman 2016; O’Kelly et al. 2015; de Sá et al. 2015). In contrast, in case of the single allocation HLP, the algorithms proposed based on Benders decomposition are quite scarce. de Camargo et al. (2011) addressed the single allocation hub location problem under congestion and proposed a hybrid algorithm combining the outer-approximation technique and a specialized version of Benders decomposition procedure for solving large-scale instances of up to 200 nodes to optimality. de Camargo et al. (2013) proposed a Benders decomposition algorithm for the many-to-many hub location routing problem. They strengthened their algorithm using Pareto optimal cuts and minimum infeasible subsystems and solved instances of up to 100 nodes to optimality. de Sá et al. (2013) considered the tree of hubs location problem where the hubs were connected by a spanning tree. They proposed an accelerated Benders decomposition algorithm to solve the problem and showed that the proposed technique is powerful enough to solve to optimality instances of up to 100 nodes.

3 Mathematical Programming Models

Let $G = (N, A)$ be a complete graph, where N is the set of nodes and A is the set of arcs such that $A \subseteq N \times N$. Assume $H \subseteq N$ be the subset of nodes that are available to locate hubs. For every $i, j \in N$, let w_{ij} denote the amount of flow originated at node i and destined to node j . Define C_{ijkm} as the transportation cost per unit flow from node i to node j routed via hubs k and m in that order:

$$C_{ijkm} = \chi c_{ik} + \alpha c_{km} + \delta c_{mj}$$

where c_{ij} denotes the transportation cost per unit flow from node i to node j which is linearly proportional to the distance between the nodes i and j and therefore satisfies the triangle inequality. χ , α , and δ are cost coefficients associated respectively with collection, transfer, and distribution links on each route. More specifically, α is the volume discount factor to reflect a scale economies on transportation cost via the inter-hub links ($0 \leq \alpha \leq 1$; $\chi \geq \alpha$; $\delta \geq \alpha$). Let the variable X_{ijkm} denote the fraction of flow w_{ij} that is sent from node i to node j using the link between the hubs k and m . Also, let the binary variable $Y_{ik} \in \{0, 1\}$ be 1 if node i is allocated to hub k and 0, otherwise. The problem then consists of selecting a set of nodes which will act as hubs, deciding on how the non-hub nodes will be allocated to these hubs, and determining the way the O/D flows will be routed within the network so that the total system-wide costs are minimized.

In the USAHLP, it is assumed that the number of hubs to be opened in the network is endogenously determined by the model, i.e., the model determines the optimal number of hubs based on the trade-off between transportation costs and fixed costs associated with opening hubs. Assuming that f_k represents the fixed cost of establishing a hub at node $k \in H$, the MIP model for the USAHLP can be written as (Skorin-Kapov et al. 1996):

(Model-USAHLP)

$$\min \sum_{k \in H} f_k Y_{kk} + \sum_{i \in N} \sum_{j \in N} \sum_{k \in H} \sum_{m \in H} C_{ijkm} w_{ij} X_{ijkm} \quad (1)$$

$$\text{s.t.:} \quad \sum_{k \in H} Y_{ik} = 1 \quad \forall i \in N \quad (2)$$

$$Y_{ik} \leq Y_{kk} \quad \forall i \in N, k \in H \quad (3)$$

$$\sum_{m \in H} X_{ijkm} = Y_{ik} \quad \forall i, j \in N, k \in H \quad (4)$$

$$\sum_{k \in H} X_{ijkm} = Y_{jm} \quad \forall i, j \in N, m \in H \quad (5)$$

$$Y_{ik} \in \{0, 1\} \quad \forall i \in N, k \in H \quad (6)$$

$$X_{ijkm} \geq 0 \quad \forall i, j \in N, k, m \in H \quad (7)$$

The objective function (1) minimizes the total transportation as well as facility location costs. Constraints (2) imply that each node i must be assigned to exactly one hub. Constraints (3) state that non-hub nodes can only be allocated to the nodes that

have already been selected as hubs. Constraints (4) state that if node i is assigned to hub k , all the outgoing flow corresponding to node i aiming to a specific node j must go through some hub m . Based on a similar reasoning, constraints (5) say that the incoming flow to a node j assigned to hub m from any node i must be transferred using some hub k . Equations 6 and 7 are the standard domain constraints for the model variables.

Unlike the USAHLP in which the number of hubs is determined by the model, in USA p HMP this number is a pre-specified quantity which is fed to the model as an input parameter. Normally, in such a case, the fixed hub establishment costs are not included in the objective function. Letting p denote the number of hub that must be located in the network, the USA p HMP can be formulated as:

(Model-USA p HMP)

$$\min \sum_{i \in N} \sum_{j \in N} \sum_{k \in H} \sum_{m \in H} C_{ijkm} w_{ij} X_{ijkm} \quad (8)$$

$$\begin{aligned} \text{s.t.:} \quad & \sum_{k \in H} Y_{kk} = p \\ & (2) - (7) \end{aligned} \quad (9)$$

The objective function (8) minimizes the total transportation costs. Constraint (9) forces the number of opened hubs to be equal to p .

4 Benders Reformulation

As previously mentioned, Benders decomposition is a classical partitioning method for MIP problems that decomposes the original problem into two simpler ones: an integer master problem (MP) and a linear subproblem (SP). The MP is a relaxed version of the original problem with the set of integer variables classified as complicating variables and its associated constraints. There is an additional continuous variable, representing a transportation cost underestimator in our specific problems. The SP is the reduced version of what would be the original problem with the values of the integer variables temporarily fixed by the MP. The algorithm solves each one of the two simpler problems iteratively, one at a time. At each iteration, a new constraint originated by the dual problem of the SP, known as Benders cut, is added to the MP. After a finite number of iterations, the algorithm converges to an optimal solution for the original MIP, if one exists.

In this section, we introduce a Benders reformulation of the models presented in Section 3 for the USAHLP and the USA p HMP. We use a modern implementation of Benders algorithm where a branch-and-cut framework is employed to solve the master problem on a single search tree utilizing recent developments in off-the-shelf solvers. This strategy is referred to as Branch-and-Benders-cut in the literature and has shown to yield promising results (Rahmaniani et al. 2017). In this method, the separated Benders cuts are added to the master problem whenever an incumbent

integer solution is found in the branch-and-cut tree. Using this strategy, the computational burden of solving an integer problem at each iteration is avoided. To further accelerate the convergence of our algorithms, a new cut disaggregation scheme is introduced. Strong optimality cuts, as well as an efficient procedure for solving the dual subproblems are also introduced to speed up the convergence of the algorithms.

4.1 Master Problem and Subproblem

By fixing the binary variables $Y_{ik} = \hat{Y}_{ik}$, in the Model-USAHLP and the Model-USAPHMP, the subproblem SP for both the UASHLP and USAPHMP can be written as:

(SP)

$$\min \sum_{i \in N} \sum_{j \in N} \sum_{k \in H} \sum_{m \in H} C_{ijkm} w_{ij} X_{ijkm} \quad (10)$$

$$\text{s.t.:} \quad \sum_{m \in H} X_{ijkm} = \hat{Y}_{ik} \quad \forall i, j \in N, k \in H \quad (11)$$

$$\sum_{k \in H} X_{ijkm} = \hat{Y}_{jm} \quad \forall i, j \in N, m \in H \quad (12)$$

$$X_{ijkm} \geq 0 \quad \forall i, j \in N, k, m \in H \quad (13)$$

Note that since for every $i \in N$ we have $\sum_{k \in H} \hat{Y}_{ik} = 1$, there exists exactly one possible path for each O/D flow and hence, the SP is feasible and has a unique feasible solution (the optimal solution). Let o_i and o_j denote the hubs to which nodes i and j are assigned, respectively. Then the optimal values of the decision variables X_{ijkm} can be determined as:

$$X_{ijkm}^* = \begin{cases} 1, & \text{if } k = o_i \text{ and } m = o_j \\ 0, & \text{otherwise} \end{cases} \quad \forall i, j \in N \quad (14)$$

Therefore, the optimal value of the objective function for the SP can be calculated as: $\sum_{i \in N} \sum_{j \in N} w_{ij} C_{ijo_i o_j}$. On the other hand, since the unit routing costs C_{ijkm} are finite and because of constraints (11) and (12), the SP is bounded. Hence, by strong duality it can be concluded that the dual of SP is also feasible and bounded. Let σ_{ijk} and π_{ijm} be the dual variables associated with constraints (11) and (12) in the SP, respectively. The dual subproblem DSP can now be written as:

(DSP)

$$\max \sum_{i \in N} \sum_{j \in N} \sum_{k \in H} \hat{Y}_{ik} \sigma_{ijk} + \sum_{i \in N} \sum_{j \in N} \sum_{m \in H} \hat{Y}_{jm} \pi_{ijm} \quad (15)$$

$$\text{s.t.:} \quad \sigma_{ijk} + \pi_{ijm} \leq C_{ijkm} w_{ij} \quad \forall i, j \in N, k, m \in H \quad (16)$$

$$\sigma_{ijk}, \pi_{ijm} \in \mathbb{R} \quad \forall i, j \in N, k, m \in H \quad (17)$$

It is worth mentioning that although the SP has a unique (optimal) solution, one cannot make such a claim for the DSP. Indeed, as it will be discussed later, the DSP has

multiple optimal solutions. Therefore, it is important to develop an efficient method for solving the DSP and choosing appropriate values for the optimal dual values.

Introducing a non-negative variable θ , the master problem for the USAHLP solves the following MIP model:

(MP-USAHLP)

$$\min \sum_{k \in H} f_k Y_{kk} + \theta \quad (18)$$

$$\text{s.t.: } \theta \geq \sum_{i \in N} \sum_{k \in H} \left(\sum_{j \in N} \sigma_{ijk}^t \right) Y_{ik} + \sum_{j \in N} \sum_{m \in H} \left(\sum_{i \in N} \pi_{ijm}^t \right) Y_{jm} \quad \forall t \in T = \{1, \dots, |T|\} \quad (19)$$

$$(2), (3), (6) \quad (19)$$

$$\theta \geq 0 \quad (20)$$

in which (π^t, σ^t) is the t th extreme point of the feasible solution space of the DSP defined by Eqs. 16–17. Observe that constraints (2) and (3) assure the installation of at least one hub in the network which in turn guarantees the feasibility of the SP at any iteration. Therefore, there is no need for adding feasibility cuts to the MP at each iteration.

Removing the fixed facility location cost term from the objective function and adding the constraint associated with locating p hubs, the MP for the USA p HMP can be written as follows:

(MP-USA p HMP)

$$\begin{aligned} \min \quad & \theta \\ \text{s.t.:} \quad & (2), (3), (6), (9), (19), (19) \end{aligned} \quad (21)$$

4.2 Disaggregating the Benders Cuts

A deeper analysis of the DSP reveals that one can decompose it into smaller problems, one for each $i - j$ pair. Exploiting this fact, we can disaggregate the Benders cut (19) resulting in the following set of Benders cuts:

$$\theta_{ij} \geq \sum_{k \in H} \sigma_{ijk}^t Y_{ik} + \sum_{m \in H} \pi_{ijm}^t Y_{jm} \quad \forall i, j \in N, t \in T \quad (22)$$

Therefore, rather than adding a single cut to the MP per iteration, we can take advantage of the special structure of the SP and add more than one constraint to the MP at a time, obtaining the following modified MPs for the USAHLP and the USA p HMP, respectively:

(MP ^{ij} -USAHLP)

$$\min \sum_{k \in H} f_k Y_{kk} + \sum_{i \in N} \sum_{j \in N} \theta_{ij} \quad (23)$$

$$\text{s.t.: } (2), (3), (6), (21) \quad (23)$$

$$\theta_{ij} \geq 0 \quad \forall i, j \in N \quad (24)$$

(MP^{*ij*}-USA_{*p*}HMP)

$$\begin{aligned} \min \quad & \sum_{i \in N} \sum_{j \in N} \theta_{ij} \\ \text{s.t.:} \quad & (2), (3), (6), (9), (21), (23) \end{aligned} \quad (25)$$

The number of added optimality cuts at each iteration of the algorithm using the above disaggregation scheme is n^2 where n is the number of nodes in the network ($n = |N|$).

Some authors have suggested a second scheme for the cut disaggregation in the Benders decomposition algorithm applied to the HLP (de Camargo et al. 2008; Contreras et al. 2011a). They argue that the reduction in the number of iterations (and hence, the reduction in the total solution time of the algorithm) as a result of adding n^2 cuts (one corresponding to each $i - j$ pair) per iteration of algorithm is not justified by the increased computational effort required for solving the relaxed master problems with n^2 cuts per iteration even for small sized instances. Therefore, they propose another cut disaggregation scheme in which the DSP is decomposed into n smaller problems each corresponding to a node in the network as in Contreras et al. (2011a). To this end, the new set of Benders cuts can be formulated as follows:

$$\theta_i \geq \sum_{j \in N | j \geq i} \sum_{k \in H} \sigma_{ijk}^t Y_{ik} + \sum_{j \in N | j \geq i} \sum_{m \in H} \pi_{ijm}^t Y_{jm} \quad \forall i \in N, t \in T \quad (26)$$

Based on the second disaggregation scheme proposed above, the master problems respectively for the USAHLP and the USA_{*p*}HMP can be written as follows:

(MP^{*i*}-USAHLP)

$$\begin{aligned} \min \quad & \sum_{k \in H} f_k Y_{kk} + \sum_{i \in N} \theta_i \\ \text{s.t.:} \quad & (2), (3), (6), (25) \end{aligned} \quad (27)$$

$$\theta_i \geq 0 \quad \forall i \in N \quad (28)$$

(MP^{*i*}-USA_{*p*}HMP)

$$\begin{aligned} \min \quad & \sum_{i \in N} \theta_i \\ \text{s.t.:} \quad & (2), (3), (6), (9), (25), (27) \end{aligned} \quad (29)$$

We proposed a third type of disaggregation for the Benders cuts based on a special property of the optimal solutions to the problems at hand. Since there are no capacity constraints on the installed hubs in our problems and also every non-hub node is assigned to a single hub, the minimum-cost path corresponding to an O/D pair $i - j$ ($i, j \in N$) is also a minimum-cost path for the pair $j - i$. In other words, the flows w_{ij} and w_{ji} are routed on the same path but in opposite directions and therefore, we can simply reduce the size of our models by just taking into account the pairs $i - j$ such that $j \geq i$. In order to get smaller formulations and also to prevent the double calculation of the transportation cost for a demand originating and ending at the same

node (i.e., if $w_{ii} \neq 0$), we define a new flow parameter D_{ijkm} for $i, j \in N$ ($j \geq i$) as follows:

$$D_{ijkm} = \begin{cases} w_{ij}C_{ijkm} + w_{ji}C_{jimk}, & \text{if } i \neq j \\ w_{ii}C_{iikm}, & \text{if } i = j. \end{cases}$$

Based on the above discussion, more compact formulation for the SP can be obtained by replacing the objective function (10) with $\sum_i \sum_{j|j \geq i} \sum_k \sum_m D_{ijkm} X_{ijkm}$, and also imposing the constraints (11)–(13) for $i, j \in N$ such that $j \geq i$. Accordingly, the required changes can be easily made to the DSP. Hence, the new disaggregated Benders cuts can now be written as:

$$\theta_{ij} \geq \sum_{k \in H} \sigma_{ijk}^t Y_{ik} + \sum_{m \in H} \pi_{ijm}^t Y_{jm} \quad \forall i, j \in N (j \geq i), t \in T \quad (30)$$

The corresponding MPs for the USAHLP and the USA p HMP can be stated as follows:

(MP $^{i \leq j}$ -USAHLP)

$$\min \sum_{k \in H} f_k Y_{kk} + \sum_{i \in N} \sum_{j \in N | j \geq i} \theta_{ij} \quad (31)$$

$$\text{s.t.: } (2), (3), (6), (29)$$

$$\theta_{ij} \geq 0 \quad \forall i, j \in N (j \geq i) \quad (32)$$

(MP $^{i \leq j}$ -USA p HMP)

$$\begin{aligned} \min \quad & \sum_{i \in N} \sum_{j \in N | j \geq i} \theta_{ij} \\ \text{s.t.: } \quad & (2), (3), (6), (9), (29), (301) \end{aligned} \quad (33)$$

Note that the number of added optimality cuts at each iteration of the algorithm using the new disaggregation scheme is $n(n+1)/2$.

4.3 Solving Dual Subproblem and Strengthening Benders Cuts

Magnanti and Wong (1981) introduced the concept of nondominated or Parteo-optimal cuts, i.e., the strong cuts that can improve the convergence of the Benders decomposition algorithm. They noticed that, the DSP has multiple optimal solutions in most applications, that results in generation of different Benders cuts. Therefore, instead of using all of these cuts, they proposed a slightly different DSP in order to identify the strongest non-dominated cuts and add them to the MP. The condition under which a cut dominates another cut is problem-specific. In case of our problem, the cut generated for the dual solution $(\sigma^{t_1}, \pi^{t_1})$ dominates the cut corresponding to the dual solution $(\sigma^{t_2}, \pi^{t_2})$ if

$$\sum_{i \in N} \sum_{k \in H} \sum_{j \in N} \sigma_{ijk}^{t_1} Y_{ik} + \sum_{j \in N} \sum_{m \in H} \sum_{i \in N} \pi_{ijm}^{t_1} Y_{jm} \geq \sum_{i \in N} \sum_{k \in H} \sum_{j \in N} \sigma_{ijk}^{t_2} Y_{ik} + \sum_{j \in N} \sum_{m \in H} \sum_{i \in N} \pi_{ijm}^{t_2} Y_{jm}$$

for all values of Y_{ik} and Y_{jm} , and with a strict inequality for at least one point. In other words, the larger the values of the dual variables σ_{ijk} and π_{ijm} the stronger the resulting Benders cut. A cut is said to be a Pareto-optimal cut if no other cut

dominates it. These strong cuts are normally generated when the DSP has alternative (multiple) optimal solutions. In this case, an optimal solution to DSP should be determined in such a manner that it provides a strong Benders cut. In order to generate strong cuts, we take a similar approach to the two-phase method introduced by Roy (1986) for the capacitated facility location problem which has been implemented on different problems in the literature (Üster et al. 2007; Üster and Aghahari 2011). We note that in case of our problems, the dual values associated with \hat{Y}_{ik} (or \hat{Y}_{jm}) with zero values do not affect the optimal objective value of the DSP. In other words, whenever $\hat{Y}_{ik} = 0$ (or $\hat{Y}_{jm} = 0$), we can modify the value of corresponding dual variable without altering the value of objective function, as long as that the new values are still feasible (i.e., constraints (16) are satisfied). Based on this short discussion, an efficient two-phase procedure for solving the DSP is proposed as follows.

Phase I In this phase, we consider a modified version of the DSP including only the dual variables σ_{ijk} and π_{ijm} corresponding to parameters \hat{Y}_{ik} and \hat{Y}_{jm} with nonzero values. It should be noted that according to constraint (2), for each node $i \in N$ there is only one hub $k \in H$ such that $\hat{Y}_{ik} = 1$. Using the notations o_i and o_j for the hubs to which the nodes i and j are assigned, the DSP for each $i - j$ pair ($i, j \in N$) can be written as:

(DSP^{*ij*})

$$\max \quad \sigma_{ijo_i} + \pi_{ijo_j} \quad (34)$$

$$\text{s.t.:} \quad \sigma_{ijk} + \pi_{ijm} \leq C_{ijkm} w_{ij} \quad \forall k, m \in H \quad (35)$$

$$\sigma_{ijk}, \pi_{ijm} \in \mathbb{R} \quad \forall k, m \in H \quad (36)$$

It is clear that the optimal objective value of the above model is $C_{ijo_i o_j} w_{ij}$ and the corresponding optimal values of variables σ_{ijo_i} and π_{ijo_j} satisfy the following equation:

$$\sigma_{ijo_i}^* + \pi_{ijo_j}^* = C_{ijo_i o_j} w_{ij} \quad (37)$$

which means that the optimal solution is not unique and there are multiple optima. For instance, given any real-valued variable λ , the optimal values of σ_{ijo_i} and π_{ijo_j} can be determined as:

$$\sigma_{ijo_i}^* = \lambda \quad (38)$$

$$\pi_{ijo_j}^* = C_{ijo_i o_j} w_{ij} - \lambda \quad (39)$$

One reasonable value for the λ which we used in our computational experiments is $\lambda = \frac{1}{2} C_{ijo_i o_j} w_{ij}$.

Phase II Having fixed the values of the dual variables σ_{ijo_i} and π_{ijo_j} as described in the first phase, we now need to determine the values of the remaining dual variables in such a way that the resulting Benders cuts be as strong as possible. To this end, for each $i - j$ pair ($i, j \in N$), one can solve the following linear programming model to

obtain the values of the dual variables corresponding to the \hat{Y}_{ik} and \hat{Y}_{jm} parameters with zero values:

$$\max \sum_{k \neq o_i} \sigma_{ijk} + \sum_{m \neq o_j} \pi_{ijm} \quad (40)$$

$$\text{s.t.: } \sigma_{ijk} + \pi_{ijm} \leq C_{ijkm} w_{ij} \quad \forall k, m \in H, k \neq o_i \text{ or } m \neq o_j \quad (41)$$

$$\sigma_{ijk}, \pi_{ijm} \in \mathbb{R} \quad \forall k, m \in H, k \neq o_i \text{ or } m \neq o_j \quad (42)$$

However, in this work we do not solve the above linear programs and instead, we propose a simple procedure to obtain solutions with large objective values in order to get relatively strong cuts with very small computational burden. Our procedure can be described as follows. Suppose that for each $i - j$ pair ($i, j \in N$), we have fixed the values to the variables σ_{ijo_i} and π_{ijo_j} to respectively $\sigma_{ijo_i}^*$ and $\pi_{ijo_j}^*$ as we did in the first phase. Then, for any $m \in H$ such that $m \neq o_j$, the largest possible value for the dual variables π_{ijm} can be determined as:

$$\pi_{ijm}^* = C_{ijo_im} w_{ij} - \sigma_{ijo_i}^* \quad (43)$$

Furthermore, by fixing the values of π_{ijm} , the largest values of the dual variable σ_{ijk} , for all $k \in H$ and $k \neq o_i$, can be calculated as:

$$\sigma_{ijk}^* = \min_{m \in H} \{C_{ijkm} w_{ij} - \pi_{ijm}^*\} \quad (44)$$

We note that, based on Eq. 37, the objective function value obtained for the dual subproblem of the pair $i - j$ (i.e., the DSP^{ij}), is equal to $C_{ijo_i o_j} w_{ij}$. Hence, the objective function value for the DSP, as the sum of the corresponding values for all the $i - j$ pairs, equals $\sum_{i \in N} \sum_{j \in N} C_{ijo_i o_j} w_{ij}$, which is identical to the optimal objective function value of the SP as we calculated earlier. Therefore, from the duality theory, it can be concluded that the solution obtained by the proposed procedure for the DSP is optimal. The above simple procedure obtains a very good solution to the problem (40)–(42) with large values for the dual variables which result in strong Benders cuts. In other words, the above procedure enables us to solve the DSP to optimality and obtain relatively strong Benders cut using a set of simple calculations and without resorting to an off-the-shelf solvers which helps to significantly reduce the total time spent by our proposed algorithms to obtain the optimal solutions. The pseudo-code for the proposed procedure for solving the DSP is presented in Algorithm 1.

Algorithm 1 Pseudo-code for the proposed solution procedure of the DSP

```

1: for all  $i \in N$  do
2:   for all  $s \in H$  do
3:     if  $\hat{Y}_{is} = 1$  then
4:        $o_i \leftarrow s$ 
5:     end if
6:   end for
7:   for all  $j \in N$  do
8:      $\sigma_{ijo_i}^* \leftarrow \frac{1}{2} C_{ijo_i o_j} w_{ij}$ 
9:      $\pi_{ijo_j}^* \leftarrow \frac{1}{2} C_{ijo_i o_j} w_{ij}$ 
10:    for all  $m \in H, m \neq o_j$  do
11:       $\pi_{ijm}^* \leftarrow C_{ijo_i m} w_{ij} - \sigma_{ijo_i}^*$ 
12:    end for
13:    for all  $k \in H, k \neq o_i$  do
14:       $\sigma_{ijk}^* \leftarrow \min\{C_{ijk m} w_{ij} - \pi_{ijm}^* : m \in H\}$ 
15:    end for
16:  end for
17: end for
18: return  $\sigma_{ijk}^*, \pi_{ijm}^*$  for all  $i, j \in N, k, m \in H$ .

```

5 Computational Experiments

In order to test the efficiency of the proposed solution algorithms, we use three commonly used data sets from the hub location literature, namely the CAB, TR, and AP data sets. The CAB data set introduced by O’Kelly (1987) is based on the airline passenger interactions between 25 US cities in 1970 evaluated by the Civil Aeronautics Board. This data set has been used by most of the hub location researchers in the literature. To solve the problem on the CAB data set, we have set the values of parameters χ and δ to 1 and the parameter α is considered at five levels: 0.2, 0.4, 0.6, 0.8, and 1.0. Since the CAB data set does not include fixed cost for opening hub facilities, the fixed cost for establishing a hub will be taken at four levels: 250, 200, 150, and 100 as in Abdinour-Helm and Venkataramanan (1998), Abdinour-Helm (1998), Topcuoglu et al. (2005), Cunha and Silva (2007), and Chen (2007).

The second data set we have used in our computational experiments is the TR data set (Tan and Kara 2007) which is based on the cargo flows between 81 cities of Turkey. Unlike the CAB data set, the flow matrix in the TR data set is not symmetrical. In most of the papers using the TR data set as benchmark, only 22 out of 81 these cities which have higher values of flow interactions are considered as candidate nodes for locating hubs ($|H| = 22$). However, in this paper we use the TR data set in two settings: *a*) considering only the 22 high traffic cities as candidate hub nodes, and *b*) considering all 81 cities as candidate hub nodes. Similar to the case of the CAB data set, the values of parameters χ and δ are set to 1 and the parameter α is

considered at five levels: 0.2, 0.4, 0.6, 0.8, and 1.0. The original fixed hub costs values in the TR data set are multiplied by scaling factor taking three different values as $FCS \in \{0.05, 0.1, 0.15\}$.

The third data set used in our experiments is the Australia Post (AP) data set that was first used by Ernst and Krishnamoorthy (1996). The AP data set is based on a postal delivery in Sydney, Australia and consists of 200 nodes representing postal districts. Similar to the TR data set, the flow matrix of the AP data set is not symmetrical. Furthermore, the AP data set includes demands originating and ending at the same node, i.e., $w_{ii} \neq 0$ for all $i \in N$. The parameters χ , α , and δ are set to 3, 0.75, and 2, respectively. There are two types of fixed hub establishment cost values and two types of hub capacity values in the AP data set, namely the *tight* and *loose* values. This results in four different classes of instances as the LT, TT, LL, and TL instances (the first letter stands for the type of the fixed cost and the second letter stands for the type of capacity). Since we do not consider the hub capacities in our work, we use the LT and TT instances for the USAHLP and the LL instances for the USA ϕ HMP.

All computational tests have been carried out on a computer with Intel(R) Core(TM) i3-3220 CPU of 3.30 GHz and 16 GB of RAM, using the Microsoft Windows 7 operating system. Also, the Benders decomposition has been implemented in JAVA using CPLEX version 12.6 to solve the master problems. The proposed Benders decomposition algorithms are implemented using the lazy constraint callback function available in CPLEX in which cuts are added to the master problem at each time an incumbent solution is found. The relative gap parameter of CPLEX is set to 0 and hence, all the obtained solutions are optimal.

5.1 Performance Analysis of the Two Cut Disaggregation Schemes

The aim of this part of the computational experiments is to compare the effectiveness of the different cut disaggregation schemes discussed in Section 4.2 on the performance of the proposed Benders decomposition algorithms. As noted earlier, it is shown in de Camargo et al. (2008, 2011a) that the second disaggregation scheme that adds n cuts per iteration is superior to the one adding n^2 cuts per iteration for the HLP. Therefore, we compare the performance of our new proposed disaggregation scheme (adding $n(n+1)/2$ cuts per iteration) to only the one that adds n cuts. To this end, the corresponding versions of the algorithms are implemented for both the USAHLP and USA ϕ HMP on several instances of the three aforementioned data sets.

The detailed results of applying the two Benders decomposition algorithms on nine problem instances from each of the CAB, TR, and AP data sets for USAHLP are reported in Table 1. The first and second columns under the heading “CAB” give the discount factor and fixed cost of opening hubs, respectively. The next columns show the CPU time (in seconds) needed to obtain an optimal solution for each problem instance using the two Benders decomposition algorithms. The columns entitled “FCS” show fixed hub cost scaling factor for the TR data set. Finally, the columns entitled “ $|N|$ ” and “Type” respectively show the number of nodes and the type of fixed cost in the corresponding problem instance from the AP data set.

As can be observed from the above results, for the USAHLP, the overall performance of the algorithm with $n(n+1)/2$ cuts in terms of the average CPU times is

Table 1 Effect of cut disaggregation on the algorithm's performance for USAHLP

CAB				TR ($ H = 81$)				AP			
# cuts per iteration				# cuts per iteration				# cuts per iteration			
		n	$n(n+1)/2$			n	$n(n+1)/2$			n	$n(n+1)/2$
α	f_k	CPU(s)	CPU(s)	α	FCS	CPU(s)	CPU(s)	$ N $	Type	CPU(s)	CPU(s)
0.4	100	0.65	0.38	0.4	0.05	47.33	42.59	50	LT	4.17	4.42
	150	0.64	0.39		0.1	42.35	48.93		TT	2.02	3.92
	200	0.57	0.47		0.15	25.45	55.45				
0.6	100	0.37	0.48	0.6	0.05	313.46	97.51	75	LT	15.79	41.55
	150	0.38	0.49		0.1	173.38	144.50		TT	6.02	15.38
	200	0.35	0.23		0.15	31.39	42.32				
0.8	100	0.72	0.80	0.8	0.05	1183.38	247.77	125	LT	127.86	411.46
	150	0.47	0.35		0.1	137.17	159.45		TT	97.59	188.05
	200	0.48	0.37		0.15	95.71	156.06				
Avg.		0.51	0.44	Avg.		227.74	110.51	Avg.		42.24	110.80

slightly better than the algorithm with n cuts. Although for some instances, the algorithm with n performs better in terms of the CPU time, but for the instances that take longer times to be solved, the algorithm with $n(n+1)/2$ cuts performs much better.

Table 2 Effect of cut disaggregation on the algorithm's performance for USA p HMP

CAB				TR ($ H = 81$)				AP			
# cuts per iteration				# cuts per iteration				# cuts per iteration			
		n	$n(n+1)/2$			n	$n(n+1)/2$			n	$n(n+1)/2$
α	p	CPU(s)	CPU(s)	α	p	CPU(s)	CPU(s)	$ N $	p	CPU(s)	CPU(s)
0.4	2	0.54	0.11	0.4	4	136.21	293.15	50	2	2.38	3.47
	3	0.66	0.33		6	123.24	57.97		3	2.19	3.08
	4	0.24	0.28		8	1531.97	323.08		4	2.75	4.83
0.6	2	0.27	0.13	0.6	4	450.27	363.33	75	4	18.28	41.66
	3	0.40	0.40		6	368.21	106.54		6	64.82	83.63
	4	0.37	0.47		8	5989.68	686.88		8	31.47	29.94
0.8	2	0.39	0.30	0.8	4	1071.40	1084.03	125	5	1855.96	1286.48
	3	0.43	0.45		6	1061.97	280.25		10	878.17	414.08
	4	0.72	0.69		8	40219.06	800.64		15	8391.66	1271.86
Avg.		0.45	0.35	Avg.		5661.33	443.99	Avg.		1249.74	348.78

The results for solving the USA p HMP on different instances of the three data sets using the two Benders decomposition algorithms are presented in Table 2. The columns entitled “ p ” show the number of opened hubs.

It can be seen that the difference between the average CPU times is more considerable for the USA p HMP as compared to that of the USAHLP. In other words, the algorithm with $n(n+1)/2$ cuts per iteration is much more efficient than the one with n cuts in terms of the CPU time for the USA p HMP. In particular, for the instances with relatively large solution times, the algorithm with $n(n+1)/2$ performs much better. Based on the above observations, we use the algorithm with $n(n+1)/2$ cuts in the rest of the computational experiments for both the USAHLP and the USA p HMP.

5.2 Numerical Results for the USAHLP

Table 3 shows the solution results obtained by applying the proposed Benders decomposition algorithm to the USAHLP on CAB data set. The first column gives the value of discount factor (α), whereas the second column shows the value of fixed cost

Table 3 Results for the USAHLP with the CAB data set

α	f_k	Opt. Obj.	CPU(s)		
			Model-USAHLP	Model-E&K	BD
0.2	100	1029.63	142.29	2.29	0.59
	150	1217.35	147.86	2.91	0.46
	200	1367.35	189.52	2.71	0.33
	250	1500.91	216.63	2.26	0.28
0.4	100	1187.52	132.51	5.16	0.38
	150	1351.70	154.67	3.83	0.39
	200	1501.63	233.58	7.79	0.47
	250	1601.63	193.92	2.87	0.15
0.6	100	1333.56	164.52	21.02	0.48
	150	1483.56	211.3	8.22	0.49
	200	1601.21	253.93	11.29	0.23
	250	1701.21	279.07	4.65	0.19
0.8	100	1458.83	169.57	32.71	0.80
	150	1594.08	226.37	29.4	0.35
	200	1690.58	221.13	15.51	0.37
	250	1740.58	206.63	1.85	0.19
1	100	1556.63	247.4	34.9	0.54
	150	1640.58	225.57	12.03	0.30
	200	1690.58	213.93	1.99	0.13
	250	1740.58	190.15	1.57	0.12
Average			201.02	10.25	0.37

for installing hubs. The optimal values of the objective function are reported in the column labeled as “Opt. Obj.”. To compare the efficiency of the proposed solution algorithm, we have solved the problem using the Model-USAHLP and also the three-indexed model proposed by Ernst and Krishnamoorthy (1996) using CPLEX. The CPU time, in seconds, needed to obtain the optimal solution of the problem using these models are presented in column labeled as “Model-USAHLP” and “Model-E&K”, respectively. The next column under the heading of “BD” gives the solution time for our proposed Benders decomposition algorithm.

Results reported in the above table indicate that the model proposed by Ernst and Krishnamoorthy (1996) is much more efficient than the model proposed by Skorin-Kapov et al. (1996) in terms of the solution time. Furthermore, it can be observed that the proposed Benders decomposition algorithm obtains the optimal solution for all the instances of the CAB data set in much smaller computational times (less than half of a second on average). It should be noted that for larger instances of the problem solving the mathematical models using off-the-shelf solvers is not practical as the time and memory requirements of such methods grows very rapidly by the size of the problem.

The results for solving the problem on the TR data set with $|H| = 22$ and $|H| = 81$ using the proposed Benders decomposition algorithm are shown in Table 4.

The first column in the above table shows the different values considered as hub fixed cost scaling factor. The results show that the Benders algorithm have obtained the optimal solutions for all instances in very short computational times. All the instances with 22 candidate hub nodes are solved in less than 130 seconds, whereas

Table 4 Results for the USAHLP with the TR data set using BD algorithm

FCS	α	$ H = 22$		$ H = 81$	
		Opt. Obj.	CPU(s)	Opt. Obj.	CPU(s)
0.05	0.2	549.96	14.10	547.57	28.76
	0.4	682.84	13.04	678.60	42.59
	0.6	808.93	43.17	803.24	97.51
	0.8	925.87	85.90	918.64	247.77
	1.0	1015.94	126.49	1015.94	975.20
0.1	0.2	683.10	11.29	681.67	31.20
	0.4	806.02	16.43	806.02	48.93
	0.6	920.27	29.88	920.27	144.50
	0.8	1007.68	56.41	1007.68	159.45
	1.0	1056.26	47.02	1056.26	54.94
0.15	0.2	772.66	17.09	765.28	32.26
	0.4	884.34	20.09	884.34	55.45
	0.6	983.63	18.03	983.63	42.32
	0.8	1067.22	21.16	1067.22	156.06
	1.0	1071.79	3.77	1071.79	14.56

the maximum solution time for the instances with all 81 nodes as candidate hub nodes is less than 1000 seconds. Another interesting observation from the above table is that in 9 out of 15 instances, the optimal set of hubs are the same for both the cases of $|H| = 22$ and $|H| = 81$. In the other 6 instances, although some of the opened hubs in the optimal solution are different, the optimal objective values differ very slightly for the two cases. It should be mentioned that this is the first time in the literature that the USAHLP is solved to optimality for the TR data set.

The results for solving the USAHLP on AP data set using the proposed Benders decomposition algorithm are shown in Table 5 for both the loose and tight fixed costs. The first column gives the size of the instance as the cardinality of the nodes set ($|N|$). For the AP data set, 14 different problem sizes ranging from 10 up to 200 nodes are considered. We also included the results for solving the problem by the row generation (RG) method proposed in Meier and Clausen (2017). As noted in Section 2, the proposed row generation method is based on the assumption that the transportation costs between hubs are proportional to the Euclidean distances between them. Since the distances in the AP data set are calculated based on the Euclidean norm, the RG method were applied to the AP instances with $|N| \geq 50$ in Meier and Clausen (2017). The solution times for the row generation method are reported in the column labeled as “RG”. Although the results reported in Meier and Clausen (2017) has been obtained by running the RG algorithm on a different computer, but the configuration of their machine is very similar to that of ours.

Table 5 Results for the USAHLP with the AP data set using BD and RG algorithms

$ N $	Loose fixed costs (LT)			Tight fixed costs (TT)		
	Opt. Obj.	CPU(s)		Opt. Obj.	CPU(s)	
		BD	RG (Meier and Clausen 2017)		BD	RG (Meier and Clausen 2017)
10	224250.05	0.69	–	263399.94	0.22	–
20	234690.96	0.36	–	271128.18	0.32	–
25	236650.63	0.48	–	295667.84	0.68	–
40	240986.23	2.10	–	293164.84	0.67	–
50	237421.99	4.42	1.06	300420.99	3.92	2.28
60	228855.08	10.52	7.13	264742.11	10.62	5.23
70	226188.20	21.76	3.15	261294.99	8.41	1.15
75	235847.50	41.55	12.29	288778.29	15.38	1.48
90	225475.48	98.92	7.32	257415.86	50.40	6.12
100	238016.28	82.42	6.91	305097.95	60.23	3.24
125	227949.00	411.46	43.30	258839.68	188.05	16.08
150	225450.10	1259.22	107.24	234778.74	478.38	26.30
175	227655.38	2044.77	188.17	247876.80	1639.21	44.55
200	233802.98	5493.47	68.18	272188.11	20292.35	1399.53

The results presented in Table 5 show that the proposed algorithm has solved all the AP instances to optimality in reasonable computational times. All the instances except the ones with $|N| = 200$, are solved within CPU time of smaller than an hour. Further, it can be seen that the instances with tight fixed cost values are in general solved in smaller computational time than their loose counterparts. The results also show that the RG algorithm is considerably faster than the BD algorithm. However, as noted earlier, the RG algorithm is designed for solving a special case of the problem, where the distances between hubs are calculated using the Euclidean norm. Nevertheless, in most of the real-world applications (e.g. road transportation), the real (travel) distances between hubs are substantially different from the corresponding Euclidean distances.

5.3 Numerical Results for the USA p HMP

As for the USAHLP, we first compare the performance of the proposed Benders decomposition algorithm with those of the Model-USA p HMP and the three-indexed formulation of Ernst and Krishnamoorthy (1996). The results for this analysis are

Table 6 Results for the USA p HMP with the CAB data set

α	p	Opt. Obj.	CPU(s)		
			Model-USA p HMP	Model-E&K	BD
0.2	2	1000.91	188.19	1.63	0.52
	3	767.35	152.31	2.99	0.45
	4	629.63	130.26	2.73	0.21
	5	538.37	117.22	2.33	0.18
0.4	2	1101.63	172.66	2.39	0.11
	3	901.70	156.94	4.26	0.33
	4	787.52	144.39	4.40	0.28
	5	707.69	130.19	4.04	0.20
0.6	2	1201.21	177.33	4.16	0.13
	3	1033.56	190.17	13.73	0.40
	4	939.21	186.93	10.62	0.47
	5	876.59	159.91	21.44	0.34
0.8	2	1294.08	204.15	7.35	0.30
	3	1158.83	186.69	27.31	0.45
	4	1087.66	208.08	38.92	0.69
	5	1034.10	206.10	42.66	0.49
1	2	1359.19	239.15	8.52	0.72
	3	1256.63	213.00	33.01	0.58
	4	1211.23	233.67	53.92	1.00
	5	1173.24	300.96	80.87	1.04
Average			184.91	18.36	0.44

reported in Table 6. The first column in this table shows the values of discount factor (α) while the second column shows the number of hubs to be opened. For the CAB data set, we have used four different values for the parameter p as 2, 3, 4, and 5. The optimal values of the objective function are reported in the column labeled as “Opt. Obj.”. The CPU times needed by CPLEX to obtain the optimal solutions using the two models are presented in column labeled as “Model-USA p HMP” and “Model-E&K”, respectively. The next column under the heading of “BD” gives the solution time of the proposed Benders decomposition algorithm.

Observe that in case of the USA p HMP, still the solution time for the Benders decomposition algorithm on the CAB data set is less than half of a second on average. Furthermore, The results show that the model proposed by Ernst and Krishnamoorthy (1996) is much more efficient than the model proposed by Skorin-Kapov et al. (1996) in terms of the solution time.

Table 7 Results for the USA p HMP with the TR data set using BD algorithm

α	p	$ H = 22$		$ H = 81$	
		Opt. Obj.	CPU (s)	Opt. Obj.	CPU (s)
0.2	2	784.84	9.57	781.32	33.81
	4	580.44	23.05	575.17	66.36
	6	454.09	10.76	448.60	27.83
	8	399.61	13.60	396.48	33.36
	10	359.72	16.12	357.70	23.97
0.4	2	860.76	12.15	850.17	41.19
	4	690.28	33.33	683.24	293.15
	6	586.30	21.85	579.38	57.97
	8	531.86	21.45	530.39	323.08
	10	494.42	20.18	493.30	320.94
0.6	2	916.90	15.80	916.69	142.99
	4	790.69	78.08	777.03	363.33
	6	699.64	36.83	691.35	106.54
	8	655.24	34.73	651.35	686.88
	10	622.19	50.77	619.08	797.82
0.8	2	961.83	31.75	961.83	180.57
	4	871.59	112.19	861.99	1084.03
	6	805.51	66.53	792.28	280.25
	8	770.82	66.84	762.10	800.64
	10	742.51	42.25	737.71	2165.90
1	2	992.72	73.84	992.72	272.75
	4	946.94	365.88	932.56	508.60
	6	904.83	175.61	883.85	400.02
	8	876.57	226.93	862.10	1845.28
	10	857.28	150.02	843.38	1476.57

Table 7 shows the results for solving the problem on the TR data set for $|H| = 22$ and $|H| = 81$ using the proposed Benders decomposition algorithm. To solve the TR data set, five different values are used for the number of opened hubs (p) as 2, 4, 6, 8, and 10.

The results shown in the above table reveals that for the TR data set with $|H| = 22$, all the instances except one of them are solved within 370 seconds of computational time. However, the solution time for the instances with $|H| = 81$ are considerably larger as the solution space of the problem has got larger in this case. Note that unlike the case of the USAHLP, for the USA p HMP only in 2 out of 25 instances the optimal set of hubs are the same for both the cases of $|H| = 22$ and $|H| = 81$. We note again that this is the first time in the literature that the USA p HMP is solved to optimality for the TR data set.

In order to solve the USA p HMP on the AP data set we have classified the instances with different sizes in three classes of small, medium, and large sized instances. The instances with 10, 20, 25, 40, and 50 nodes are classified as small sized instances for which the values for the parameter p are taken as $p \in \{2, 3, 4, 5\}$. The instances with 60, 70, 75, and 90 nodes are classified as medium sized instances for which the values for the parameter p are selected as $p \in \{2, 3, 4, 5, 10\}$. Finally, the instances with 100, 125, 150, 175, and 200 nodes are classified as large sized instances and the four values 5, 10, 15, and 20 are used as the parameter p for these instances. Results obtained by solving the USA p HMP for the AP data set using the proposed Benders decomposition algorithm are presented in Table 8. We also included the CPU times of solving the same instances by the row generation algorithm presented in Meier et al. (2016).

It can be observed from Table 8 that the proposed Benders algorithm solves all the small sized instances of the AP data set within 5 seconds of CPU time, whereas for the medium sized instances, the maximum solution time is less than 300 second. For the large sized instances, on the other hand, the solution times are considerably larger but only in one of the instances the solution time has exceeded 10 hours. Similar to the case of the USAHLP, the RG algorithm has much smaller CPU times than the BD algorithm. However, as noted earlier, unlike the RG algorithm which is only applicable to the problem instances where the transportation costs between the hubs are proportional to their Euclidean distances, the BD algorithm does not have such a limitation and hence, can be applied to any realistic problem instance like the TR instances.

In order to evaluate the convergence speed of the proposed solution algorithms we plotted the relative gap percent value over time (as reported in CPLEX log) for two instances of the USA p HMP with $|N| = 200$ and the number of installed hubs as $p = 5$ and 10, respectively. The plots are shown in Fig. 2. As can be seen from the gap progress plots, CPLEX finds the first feasible solution with a gap less than 5% in short time after it starts running. It is also observed that, when the gap drops below 1%, the convergence process gets slower and hence, a large amount of time is spent when the gaps is below 1%. In other words, if one decides to accept solutions within 1% of optimality, then the solution times will be substantially smaller than the reported values.

Table 8 Results for the USA_pHMP with the AP data set using BD and RG algorithms

Small sized instances				Medium sized instances				Large sized instances								
N	p	Opt. Obj.	CPU (s)	N	p	Opt. Obj.	CPU (s)	N	p	Opt. Obj.	CPU (s)					
			BD				RG (Meier et al. 2016)				BD	RG (Meier et al. 2016)	BD	RG (Meier et al. 2016)		
10	2	167493.06	0.16	–	60	2	179920.21	7.10	3.21	100	5	136929.44	313.80	356.39		
	3	136008.13	0.07	–		3	160338.58	13.45	4.96		10	106469.57	109.18	32.22		
	4	112396.07	0.05	–		4	144719.69	11.37	4.39		15	90533.52	144.40	85.95		
	5	91105.37	0.05	–		5	132850.29	14.77	6.89		20	80270.96	61.47	33.54		
	20	2	172816.69	0.11		–	10	102939.87	8.44		–	125	5	137175.68	1286.48	1104.31
20	3	151533.08	0.35	–	70	2	180093.19	7.91	4.35	10	107092.09	414.08	184.62			
	4	135624.88	0.14	–		3	160933.23	17.27	5.00		15	91494.56	1271.86	465.08		
	5	123130.09	0.18	–		4	145619.65	25.71	7.94		20	81471.65	213.76	111.03		
	25	2	175541.98	0.26		–	5	135835.20	36.01		45.34	150	5	137425.90	2989.83	1474.76
	3	155256.32	0.29	–		10	105357.99	16.17	–		10	107478.12	1148.01	412.53		
40	4	139197.17	0.34	–	75	2	180118.91	7.25	5.24	15	92050.58	1695.15	405.57			
	5	123574.29	0.24	–		3	161056.74	31.70	6.76		20	82229.39	531.99	185.34		
	2	177471.67	0.64	–		4	145734.20	41.66	11.08		175	5	139354.51	31347.15	10699.58	
	3	158830.54	2.15	–		5	136011.35	59.15	48.63		10	109744.35	10551.64	3023.02		
	4	143968.88	1.63	–		10	106364.90	42.23	–		15	94123.66	19602.93	8143.73		
50	5	134264.97	1.89	–	90	2	179821.64	42.98	12.87	20	83843.59	1778.11	271.03			
	2	178484.29	3.47	3.22		3	160437.43	46.11	20.27		5	140062.65	127546.79	17628.38		
	3	158569.93	3.08	1.86		4	145133.69	78.73	27.76		10	110147.66	46706.90	4957.31		
	4	143378.05	4.83	1.90		5	135808.25	299.83	234.17		15	94459.20	26640.56	1107.48		
	5	132366.95	4.48	2.26		10	105506.44	46.50	–		20	84955.36	27224.48	526.08		

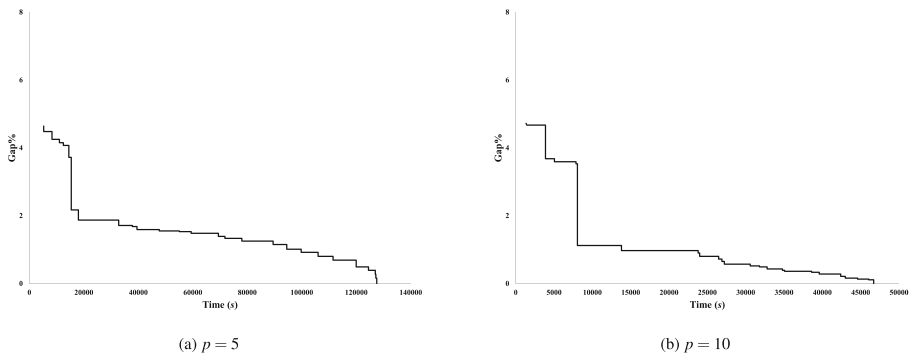


Fig. 2 Gap% change over time for the USA p HMP instances with $|N| = 200$

6 Conclusions

In this paper, we proposed efficient Benders decomposition algorithms for solving the uncapacitated single allocation hub location problem (USAHLP) and the uncapacitated single allocation p -hub median problem (USA p HMP). The standard implementation of the Benders algorithms was enhanced by applying a number of refinement techniques such as using a new cut disaggregation scheme, generating strong optimality cuts, and an efficient algorithm to solve the dual subproblems without calling the off-the-shelf optimization packages. Furthermore, a modern implementation of the algorithm was used to solve the problem on a single search tree in which the Benders cuts were successively added within a branch-and-cut framework. Computational experiments were conducted on three well-known data sets in the literature, namely the CAB, TR, and AP data sets. The instances for the TR data set had not been solved to optimality in the literature before. The proposed algorithms have shown to be able to solve all the tested instances in short and reasonable CPU times. However, the results show that solving the USA p HMP instances generally take more time than the USAHLP instances of the same size.

An interesting avenue for further research is to modify and apply the proposed algorithms to solve other variants of the single allocation HLPs such the single allocation p -hub center problem and the single allocation p -hub maximal covering problem.

Acknowledgments The authors sincerely thank the two anonymous referees for their valuable and constructive comments that helped us improve the quality and presentation of the paper.

References

- Abdinnour-Helm S (1998) A hybrid heuristic for the uncapacitated hub location problem. *Eur J Oper Res* 106(2):489–499
- Abdinnour-Helm S, Venkataramanan M (1998) Solution approaches to hub location problems. *Ann Oper Res* 78:31–50

- Abyazi-Sani R, Ghanbari R (2016) An efficient tabu search for solving the uncapacitated single allocation hub location problem. *Comput Indus Eng* 93:99–109
- Alumur S, Kara BY (2008) Network hub location problems: the state of the art. *Europ J Oper Res* 190(1):1–21
- Aykin T (1994) Lagrangian relaxation based approaches to capacitated hub-and-spoke network design problem. *Eur J Oper Res* 79(3):501–523
- Benders JF (1962) Partitioning procedures for solving mixedvariables programming problems. *Numerisch Mathematik* 4(1):238–252
- Campbell JF (1994) Integer programming formulations of discrete hub location problems. *Eur J Oper Res* 72(2):387–405
- Campbell JF, O’Kelly ME (2012) Twenty-five years of hub location research. *Transp Sci* 46(2):153–169
- Chen JF (2007) A hybrid heuristic for the uncapacitated single allocation hub location problem. *Omega* 35(2):211–220
- Contreras I (2015) Hub location problems. In: Laporte G, Nickel S, Saldanha da Gama F (eds) *Location Science*. Springer International Publishing, pp 311–344
- Contreras I, Díaz JA, Fernández E (2009) Lagrangean relaxation for the capacitated hub location problem with single assignment. *OR Spectr* 31(3):483–505
- Contreras I, Cordeau JF, Laporte G (2011a) Benders decomposition for large-scale uncapacitated hub location. *Oper Res* 59(6):1477–1490
- Contreras I, Cordeau JF, Laporte G (2011b) Stochastic uncapacitated hub location problem. *Eur J Oper Res* 212(3):518–528
- Contreras I, Díaz JA, Fernández E (2011c) Branch and price for large-scale capacitated hub location problems with single assignment. *INFORMS J Comput* 23(1):41–55
- Contreras I, Cordeau JF, Laporte G (2012) Exact solution of large-scale hub location problems with multiple capacity levels. *Transp Sci* 46(4):439–459
- Cunha CB, Silva MR (2007) A genetic algorithm for the problem of configuring a hub-and-spoke network for a ltl trucking company in brazil. *Eur J Oper Res* 179(3):747–758
- de Camargo RS, de Miranda GJr, Luna HP (2008) Benders decomposition for the uncapacitated multiple allocation hub location problem. *Comput Oper Res* 35(4):1047–1064
- de Camargo RS, de Miranda GJr, Ferreira RPM, Luna HP (2009a) Multiple allocation hub-and-spoke network design under hub congestion. *Comput Oper Res* 36(12):3097–3106
- de Camargo RS, de Miranda GJr, Luna HP (2009b) Benders decomposition for hub location problems with economies of scale. *Transp Sci* 43(1):86–97
- de Camargo RS, de Miranda GJr, Ferreira RP (2011) A hybrid outer-approximation/benders decomposition algorithm for the single allocation hub location problem under congestion. *Oper Res Lett* 39(5):329–337
- de Camargo RS, de Miranda GJr, Løkketangen A (2013) A new formulation and an exact approach for the many-to-many hub location-routing problem. *Appl Math Modell* 37(12–13):7465–7480
- de Sá EM, de Camargo RS, de Miranda GJr (2013) An improved benders decomposition algorithm for the tree of hubs location problem. *Eur J Oper Res* 226(2):185–202
- de Sá EM, Contreras I, Cordeau JF, de Camargo RS, de Miranda GJr (2015) The hub line location problem. *Transp Sci* 49(3):500–518
- Ernst AT, Krishnamoorthy M (1996) Efficient algorithms for the uncapacitated single allocation p -hub median problem. *Locat Sci* 4(3):139–154
- Ernst AT, Krishnamoorthy M (1998) An exact solution approach based on shortest-paths for p -hub median problems. *INFORMS J Comput* 10(2):149–162
- Ernst A, Krishnamoorthy M (1999) Solution algorithms for the capacitated single allocation hub location problem. *Ann Oper Res* 86(0):141–159
- Farahani RZ, Hekmatfar M, Arabani AB, Nikbakhsh E (2013) Hub location problems: a review of models, classification, solution techniques, and applications. *Comput Indus Eng* 64(4):1096–1109
- Fazel-Zarandi MM, Berman O, Beck JC (2013) Solving a stochastic facility location/fleet management problem with logic-based benders’ decomposition. *IIE Trans* 45(8):896–911
- Gelareh S, Nickel S (2008) A benders decomposition for hub location problems arising in public transport. Springer, Berlin, pp 129–134
- Gelareh S, Nickel S (2011) Hub location problems in transportation networks. *Transp Res E* 47(6):1092–1111
- Geoffrion AM, Graves GW (1974) Multicommodity distribution system design by benders decomposition. *Manag Sci* 20(5):822–844

- He X, Zheng H, Peeta S, Li Y (2018) Network design model to integrate shelter assignment with contraflow operations in emergency evacuation planning. *Netw Spatial Econ*, 1–24. <https://doi.org/10.1007/s11067-017-9381-y>
- Ilić A, Urošević D, Brimberg J, Mladenović N (2010) A general variable neighborhood search for solving the uncapacitated single allocation p -hub median problem. *Eur J Oper Res* 206(2):289–300
- Kara BY (1999) Modeling and analysis of issues in hub location problem. Bilkent University, PhD thesis
- Klineciewicz JG (1992) Avoiding local optima in the p -hub location problem using tabu search and GRASP. *Ann Oper Res* 40(1):283–302
- Kratka J, Stanimirović Z, Tošić D, Filipović V (2007) Two genetic algorithms for solving the uncapacitated single allocation p -hub median problem. *Eur J Oper Res* 182(1):15–28
- Magnanti TL, Wong RT (1981) Accelerating benders decomposition: algorithmic enhancement and model selection criteria. *Oper Res* 29(3):464–484
- Marić M, Stanimirović Z, Stanojević P (2013) An efficient memetic algorithm for the uncapacitated single allocation hub location problem. *Soft Comput* 17(3):445–466
- Meier JF, Clausen U (2017) Solving single allocation hub location problems on Euclidean data. *Transp Sci.* <https://doi.org/10.1287/trsc.2017.0751>
- Meier JF, Clausen U, Rostami B, Buchheim C (2016) A compact linearisation of euclidean single allocation hub location problems. *Electron Notes Discret Math* 52:37–44
- Merakli M, Yaman H (2016) Robust intermodal hub location under polyhedral demand uncertainty. *Transp Res B Methodol* 86:66–85
- O’Kelly ME (1987) A quadratic integer program for the location of interacting hub facilities. *Eur J Oper Res* 32(3):393–404
- O’Kelly ME (1992) Hub facility location with fixed costs. *Pap Reg Sci* 71(3):293–306
- O’Kelly ME, Luna HPL, de Camargo RS, de Miranda GJr (2015) Hub location problems with price sensitive demands. *Netw Spatial Econ* 15(4):917–945
- Peker M, Kara BY, Campbell JF, Alumur SA (2016) Spatial analysis of single allocation hub location problems. *Netw Spatial Econ* 16(4):1075–1101
- Rahmaniani R, Crainic TG, Gendreau M, Rei W (2017) The benders decomposition algorithm: a literature review. *Eur J Oper Res* 259(3):801–817
- Roy TJV (1986) A cross decomposition algorithm for capacitated facility location. *Oper Res* 34(1):145–163
- Silva MR, Cunha CB (2009) New simple and efficient heuristics for the uncapacitated single allocation hub location problem. *Comput Oper Res* 36(12):3152–3165
- Skorin-Kapov D, Skorin-Kapov J (1994) On tabu search for the location of interacting hub facilities. *Eur J Oper Res* 73(3):502–509
- 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
- Tan PZ, Kara BY (2007) A hub covering model for cargo delivery systems. *Networks* 49(1):28–39
- Taner MR, Kara BY (2016) Endogenous effects of hubbing on flow intensities. *Netw Spatial Econ* 16(4):1151–1158
- Tang L, Jiang W, Saharidis GKD (2013) An improved benders decomposition algorithm for the logistics facility location problem with capacity expansions. *Ann Oper Res* 210(1):165–190
- Tang L, Sun D, Liu J (2016) Integrated storage space allocation and ship scheduling problem in bulk cargo terminals. *IIIE Trans* 48(5):428–439
- Topcuoglu H, Corut F, Ermis M, Yilmaz G (2005) Solving the uncapacitated hub location problem using genetic algorithms. *Comput Oper Res* 32(4):967–984
- Üster H, Agrahari H (2011) A benders decomposition approach for a distribution network design problem with consolidation and capacity considerations. *Oper Res Lett* 39(2):138–143
- Üster H, Easwaran G, Akçali E, Çetinkaya S (2007) Benders decomposition with alternative multiple cuts for a multi-product closed-loop supply chain network design model. *Naval Res Logist (NRL)* 54(8):890–907
- Wentges P (1996) Accelerating benders’ decomposition for the capacitated facility location problem. *Math Methods Oper Res* 44(2):267–290
- Wolf S, Merz P (2007) Evolutionary local search for the super-peer selection problem and the p -hub median problem. In: *Hybrid Metaheuristics*. Springer, pp 1–15