



Ordered median hub location problems with capacity constraints



J. Puerto ^{a,*}, A.B. Ramos ^a, A.M. Rodríguez-Chía ^b, M.C. Sánchez-Gil ^b

^a Facultad de Matemáticas, Universidad de Sevilla, Spain

^b Facultad de Ciencias, Universidad de Cádiz, Spain

ARTICLE INFO

Article history:

Received 12 May 2014

Received in revised form 24 March 2015

Accepted 13 May 2015

Available online 15 June 2015

Keywords:

Hub location

Discrete ordered median problems

Capacities

ABSTRACT

The Single Allocation Ordered Median Hub Location problem is a recent hub model introduced in Puerto et al. (2011) that provides a unifying analysis of a wide class of hub location models. In this paper, we deal with the capacitated version of this problem, presenting two formulations as well as some preprocessing phases for fixing variables. In addition, a strengthening of one of these formulations is also studied through the use of some families of valid inequalities. A battery of test problems with data taken from the AP library are solved where it is shown that the running times have been significantly reduced with the improvements presented in the paper.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Network design problems are among the most interesting models in combinatorial optimization. In the last years researchers have devoted a lot of attention to a particular member within this family, namely the hub location problem, that combines network design and location aspects of supply chain models, see the surveys (Alumur and Kara, 2008; Campbell et al., 2002; Campbell and O'Kelly, 2012). The main advantage of using hubs in distribution problems is that they allow to consolidate shipments in order to reduce transportation costs by applying economies of scale, which are naturally incorporated to the models through discount factors. Hub location problems have been studied from different perspectives giving rise to a number of papers considering alternative criteria to be optimized: the minimization of the overall transportation cost (sum) (see Campbell, 1996; Cánovas et al., 2007; Ernst and Krishnamoorthy, 1999; García et al., 2012; Labbé et al., 2005; Marín, 2005a,b; Marín et al., 2006), the minimization of the largest transportation cost or the coverage cost (Bollapragada et al., 2006; Campbell et al., 2007; Kara and Tansel, 2000; Kara and Tansel, 2003; Kratica and Stanimirovic, 2006; Meyer et al., 2009; Tan and Kara, 2007; Wagner, 2008), etc.

Apart from the choice of the optimization criterion, another crucial aspect in the literature on hub location, and in general in any location problem, is the existence or not of capacity constraints. One can recognize that although capacitated models are more realistic, the difficulty to solve them also increases in orders of magnitude with respect to their uncapacitated counterpart. In many cases new formulations are needed and a more specialized analysis is often required to solve even smaller sizes than those previously addressed for the uncapacitated versions of the problems. For this reason, capacitated versions of hub location problems have attracted the interest of locators in the last years, see (Aykin, 1994; Campbell, 1994; Contreras et al., 2009; Correia et al., 2010a,b; Ebery et al., 2000; Boland et al., 2004; Ernst and Krishnamoorthy, 1999; Marín, 2005a).

* Corresponding author.

Along the same line, it is worth mentioning some other references that assume congestion at hubs, as congestion acts as a limit on capacity, see (De Camargo et al., 2009; Elhedhli and Hu, 2004; Marianov and Serra, 2003).

An interesting version of hub location model is the Capacitated Hub Location Problem with Single Allocation (CSA-HLP), see (Contreras et al., 2009; Correia et al., 2010a; Ernst and Krishnamoorthy, 1999). In this context, single allocation means that incoming and outgoing flow of each site must be shipped via the same hub. In contrast to single allocation models, where binary variables are required in the allocation phase, multiple allocation allows different delivery patterns which in turns implies the use of continuous variables, simplifying the problems. The CSA-HLP model incorporates capacity constraints on the incoming flow at the hubs coming from origin sites or even simpler, on the number of non-hub nodes assigned to each hub. The inclusion of capacity constraints makes these models challenging from a theoretical point of view. Regarding its applicability we cite one example described in Ernst and Krishnamoorthy (1999) based on a postal delivery application, where a set of n postal districts (corresponding to postcode districts represented by nodes) exchange daily mail. The mail between all the pairs of nodes must be routed via one or at most two mail consolidation centers (hubs). In order to meet time constraints, only a limited amount of mail could be sorted at each sorting center (mail is just sorted once, when it arrives to the first hub from origin sites). Hence, there are capacity restrictions on the incoming mail that must be sorted. The problem requires to choose the number and location of hubs, as well as to determine the distribution pattern of the mail.

The CSA-HLP has received less attention in the literature than its uncapacitated counterpart. Campbell (1994) presented the first Integer Mathematical Programming formulation for the Capacitated Hub Location Problem. This formulation was strengthened by Skorin-Kapov et al. (1996). Ernst and Krishnamoorthy (1999), proposed a new model involving three-index continuous variables and developed a solution approach based on Simulated Annealing where the bounds obtained were embedded in a branch-and-bound procedure devised for solving the problem optimally. Recently, Correia et al. (2010a) have shown that this formulation may be incomplete and they propose an additional set of inequalities to assure the validity of the model in all situations. A new formulation using only two-index variables was proposed by Labbé et al. (2005), where a polyhedral analysis and new valid inequalities were addressed. Although this formulation had only a quadratic number of variables, the number of constraints was exponential, and to solve it, the authors developed a branch-and-cut algorithm based on their polyhedral analysis. Contreras et al. (2009) presented for the same problem a Lagrangian relaxation enhanced with reduction tests that allowed the computation of tight upper and lower bounds for a large set of instances.

In two recent papers, (Puerto et al., 2011, 2013), a new model of hub location, namely the Single Allocation Ordered Median Hub Location problem (SA-OMHLP), has been introduced and analyzed. This problem can be seen as a powerful tool from a modeling point of view since it allows a common framework to represent many of the previously considered criteria in the literature of hub location. Moreover, this approach is a natural way to represent the differentiation of the roles played by the different parties (origins, hubs and destinations) in logistics networks (Fonseca et al., 2010; Kalcsics et al., 2010a,b; Marín et al., 2009). This model does not assume, in advance, any particular structure on the network (Contreras et al., 2009, 2011). Instead of that this structure is derived from the choice of the parameters defining the objective function. Apart from the above mentioned characteristics, ordered median objectives are also useful to obtain robust solutions in hub problems by applying k -centrum, trimmed-mean or anti-trimmed-mean criteria. It is worth mentioning that although it is called single allocation, its meaning slightly differs from the classical interpretation in hub location where each site is allocated to just one hub and all the incoming and outgoing flow to/from this site is shipped via the same link (the one joining this site and its allocated hub). In this model, single allocation means that all the outgoing flow is delivered through the same hub, but the incoming flow can come from different hubs. Actually, this is a mixed model and basically the same situation described above, about postal deliveries, naturally fits in this framework assuming that letters from the same origin should be sorted, with respect to their destinations, in the same place and from there they are delivered via their cheapest routes. Observe that in this scheme it is also natural that incoming flow in a final destination comes from different hubs.

The SA-OMHLP distinguishes among segmented origin–destination deliveries giving different scaling factors to the origin–hub, hub–hub and hub–destination links. The cost of each origin–first hub link is scaled by a factor that depends on the position of this cost in the ordered sequence of costs from each origin to its corresponding first hub (Boland et al., 2006; Marín et al., 2009; Nickel and Puerto, 2005). Moreover, the overall interhub cost and hub–destination cost are multiplied by different economies of scale factors. The goal is to minimize the overall shipping cost under the above weighting scheme. The reader may note that the first type of scaling factors mentioned above adds a “sorting” problem to the underlying hub location model, making its formulation and solution much more challenging. This model and two different formulations were introduced in Puerto et al. (2013) while a specialized B&B&Cut algorithm was developed in Puerto et al. (2011) and Ramos (2012). None of those formulations could handle capacities since the computation burden of the problems was very high demanding. Thus, the SA-OMHLP with capacity constraints, i.e. Capacitated Single Allocation Ordered Median Hub Location problem (CSA-OMHLP) is currently an open line of research.

In this paper, we analyze in depth the CSA-OMHLP trying to obtain a better knowledge and alternative ways to solve it. Thus, the contributions of this paper are threefold. First, it combines for the first time three challenging elements in location analysis: hub facilities, capacities and ordered median objectives, proposing a promising IP formulation which remarkably reduces the number of decision variables. Second, we develop new theoretical results which provide additional knowledge on the structure of the solutions set and allow to design an efficient cut and branch approach to solve the problem. Indeed, we are giving a better polyhedral description of the considered problem (fourteen families of valid inequalities). In addition to this strengthening, we also develop two new preprocessing phases which are shown to be very effective in solving the

problem. Finally, despite the difficulty of considering simultaneously capacitated models, hubs and ordering, the techniques proposed in this paper are not only a void theoretical development but also of practical importance: this is evidenced by the improvement in more than 88% in the CPU time for solving the test problems in the paper.

The paper is organized as follows: in Section 2 we will provide, first, a MIP formulation for the capacitated version of the problem extending the one in Puerto et al. (2011) and then another formulation in the spirit of Puerto et al. (2013) where the number of variables has been considerably reduced with respect to the previous one. Section 3 strengthens the latter formulations with variable fixing and several new families of valid inequalities. In Section 4, the effectiveness of the proposed methodology is tested with an extensive computational experience comparing the performance of the two formulations and the proposed strengthening. Finally, the paper ends with some conclusions.

2. Model and MIP formulations

The goal of this paper is to analyze the CSA-OMHLP. For this reason, we elaborate from the most promising formulations of the non-capacitated version of that problem, namely the so called radius (covering) formulations, see (Puerto et al., 2011, 2013). In order to be self-contained and for the sake of readability, we include next a concise description of these formulations in their application to the capacitated problem.

Let $A = \{1, \dots, n\}$ be a set of n client sites, where each site is collecting or gathering some commodity that must be sent to the remaining ones. It is assumed, without loss of generality, that the set of candidate sites for establishing hubs is also A . Let $w_{jm} \geq 0$ be the amount of commodity to be supplied from the j -th to the m -th site for all $j, m \in A$, and let $W_j = \sum_{m \in A} w_{jm}$ be the total amount of commodity to be sent from the j -th site. Let $c_{jm} \geq 0$ denote the unit cost of sending commodity from site j to site m (not necessarily satisfying the triangular inequality). It is assumed free self-service, i.e., $c_{jj} = 0$, $\forall j \in A$.

We assume that the number of hubs to be located is $p \leq n$ and the flow pattern between each pair of sites (links connecting pairs) traverses at least one and no more than two hubs from X . Moreover, considering b_j as the capacity of a hub located at site j with $j \in A$, a solution for the problem is a set of sites $X \subseteq A$ with $|X| = p$ and enough capacity to cover the flow coming from the sites; plus a set of links connecting pairs (flow patterns) of sites j, m for all $j, m \in A$.

As it was mentioned in Section 1, the main advantage of using hubs is to reduce costs by applying economies of scale to consolidated flows in some part of the network. In this model the transportation cost is decoupled into the three differentiated possible links: origin site–first hub, hub-to-hub, and hubs–final destination. These transportation costs are scaled in a different way. The model weights origin site–first hub transportation costs by using parameters $\lambda = (\lambda_1, \dots, \lambda_n)$, with $\lambda_i \geq 0 \forall i \in A$, depending on their ordered rank values. This is, let \hat{c}_{jk} be the cost of the overall flow sent from the origin site j if it were delivered via the first hub k , i.e. $\hat{c}_{jk} := c_{jk}W_j$, $j, k \in A$. If \hat{c}_{jk} were ranked in the i -th position among all these costs, then this term would be scaled by λ_i in the objective function. For the two remaining links there are two compensation factors: $0 < \mu < 1$ for the deliveries between hubs, and $0 < \delta < 1$, $\mu < \delta$, for the deliveries between hubs and final destination sites. These parameters imply that, even in the case where the costs satisfy the triangle inequality, using a second hub could result in a cheaper connection than going directly from the first hub to the final destination. Actually, it represents the application of the economies of scale by the consolidation of flow in the hubs.

In the following we present a first valid formulation of the CSA-OMHLP, based on covering variables (the reader is referred to Puerto et al. (2011) and Puerto et al. (2013) for further details). Sorting the different delivery costs values (\hat{c}_{jk}) for $j, k \in A$, in increasing order, we get the ordered cost sequence:

$$\hat{c}_{(1)} := 0 < \hat{c}_{(2)} < \dots < \hat{c}_{(G)} := \max_{1 \leq j, k \leq n} \{\hat{c}_{jk}\}.$$

where G is the number of different elements of the above cost sequence. For convenience we consider $\hat{c}_{(0)} := 0$.

For $i \in A$ and $h = 1, \dots, G$, we define the following set of covering variables,

$$\bar{u}_{ih} := \begin{cases} 1, & \text{if the } i\text{-th smallest allocation cost is at least } \hat{c}_{(h)}, \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

Clearly, the i -th smallest allocation cost is equal to $\hat{c}_{(h)}$ if and only if $\bar{u}_{ih} = 1$ and $\bar{u}_{i, h+1} = 0$.

In addition, this formulation uses two more sets of variables:

$$x_{jk} = \begin{cases} 1, & \text{if the commodity sent from origin site } j \text{ goes first to the hub } k, \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

$s_{k\ell m}$ = flow that goes through a first hub k and a second hub ℓ with destination m ,

with $j, k, \ell, m \in A$. Since we assume that any origin is allocated to itself if it is a hub, the above definition implies that site k is opened as a hub if the corresponding variable x_{kk} takes the value 1.

The formulation of the model is:

$$(F_u) \quad \min \sum_{i \in A} \sum_{h=2}^G \lambda_i (\hat{c}_{(h)} - \hat{c}_{(h-1)}) \bar{u}_{ih} + \sum_{k \in A} \sum_{\ell \in A} \sum_{m \in A} (\mu c_{k\ell} + \delta c_{\ell m}) s_{k\ell m} \quad (3)$$

$$\text{s.t.} \quad \sum_{k \in A} x_{jk} = 1, \quad \forall j \in A \quad (4)$$

$$\sum_{j \in A} x_{jk} \leq n x_{kk}, \quad \forall k \in A \quad (5)$$

$$\sum_{k \in A} x_{kk} = p \quad (6)$$

$$\sum_{\ell \in A} s_{k\ell m} = \sum_{j \in A} w_{jm} x_{jk}, \quad \forall k, m \in A \quad (7)$$

$$s_{k\ell m} \leq \sum_{j \in A} w_{jm} (1 - x_{mm}) \quad \forall k, \ell, m \in A, \quad \ell \neq m \quad (8)$$

$$\sum_{\ell \in A} \sum_{m \in A} s_{k\ell m} \leq x_{kk} \sum_{j \in A} w_j, \quad \forall k \in A \quad (9)$$

$$\sum_{k \in A} \sum_{m \in A} s_{k\ell m} \leq x_{\ell\ell} \sum_{j \in A} w_j, \quad \forall \ell \in A \quad (10)$$

$$\sum_{j \in A} w_j x_{jk} \leq b_k x_{kk}, \quad \forall k \in A \quad (11)$$

$$\sum_{i \in A} \bar{u}_{ih} = \sum_{j \in A} \sum_{\substack{k=1 \\ \hat{c}_{jk} \geq \hat{c}_{(h)}}}^n x_{jk}, \quad \forall h = 1, \dots, G \quad (12)$$

$$\bar{u}_{ih} \geq \bar{u}_{i-1,h}, \quad \forall i \in A \setminus \{1\}, h = 1, \dots, G \quad (13)$$

$$\bar{u}_{ih}, x_{jk} \in \{0, 1\}, s_{k\ell m} \geq 0, \quad \forall i, j, k, \ell, m \in A, \quad h = 1, \dots, G \quad (14)$$

The objective function (3) accounts for the weighted sum of the three components of the shipping cost, namely origin–first hub, hub–hub and hub–destination. The origin–hub costs are accounted after assigning the lambda parameters, i.e. $\sum_{i \in A} \sum_{h=2}^G \lambda_i \cdot (\hat{c}_{(h)} - \hat{c}_{(h-1)}) \cdot \bar{u}_{ih}$. In addition, the second and third blocks of delivery costs, i.e. the hub-hub and hub–destination cost, scaled with the μ and δ parameters respectively, can be stated as: $\sum_{k \in A} \sum_{\ell \in A} \sum_{m \in A} (\mu c_{k\ell} + \delta c_{\ell m}) s_{k\ell m}$.

Constraints (4) ensure that the flow from the origin site j is associated with a unique first hub. Constraints (5) ensure that any origin only can be allocated to an open hub. Constraint (6) fixes the number of hubs to be located. Constraints (7) are flow conservation constraints, such that the flow that enters any hub k with final destination m is the same as the flow that leaves hub k with destination m . Constraints (8) ensure that if the final destination site is a hub, then the flow goes at most through one additional hub. These constraints are redundant whenever the cost structure satisfies the triangular inequality, however they are useful in reducing solution times (see Puerto et al., 2011). Constraints (9) and (10) establish again that the intermediate nodes in any origin–destination path should be open hubs. Constraints (11) establish the capacity constraints of the hubs. Observe that this family of constraints make redundant the family (5), but we have kept it because it reduces the computational times. Constraints (12) link sorting and covering variables. They state that the number of allocations with a cost at least $\hat{c}_{(h)}$ must be equal to the number of sites that support shipping costs to the first hub greater than or equal to $\hat{c}_{(h)}$. Finally, constraints (13) are a group of sorting conditions on the variables \bar{u}_{ih} .

The reader may note that this formulation is a natural extension for the capacitated version of the radius formulation already considered for the uncapacitated ordered median hub location problem in Puerto et al. (2011, 2013). However, although this formulation is enough to specify the CSA-OMHLP, we have found that for solving medium sized problems it produces very large MIP models, which are difficult to solve with standard MIP solvers (CPLEX, Xpress, Gurobi...). Therefore, some alternatives should be investigated.

One way to improve the performance of the above formulation is to take advantage of some features of that model to reduce the number of variables. In this case, one can succeed reducing the number of u variables. The logic of the above formulation can be further strengthened for important particular cases of the discrete ordered median hub location problem. In the following, we show a reformulation that is based on taking advantage of sequences of repetitions in the λ -vector. (See Marín et al., 2010; Puerto et al., 2013; Ramos, 2012 for similar reformulations applied to other location problems.)

One can realize that for λ -vectors with sequences of repetitions – i.e. the center, k -centrum, trimmed means or median among others, many variables used in formulation F_u are not necessary (since they are multiplied by zero in the objective function), and some others can be glued together (since they have the same coefficient in the objective function). Moreover, under the assumption of the free self-service, and that any origin is allocated to itself if it is a hub, we conclude that the p smallest transportation costs from the origin to the first hubs are 0, i.e. the first p components of the λ -vector are multiplied by 0. Therefore, in order to simplify the problem one can disregard the p first components of the λ -vector. Let $\tilde{\lambda} = (\tilde{\lambda}_1, \dots, \tilde{\lambda}_{n-p}) := (\lambda_{p+1}, \dots, \lambda_n)$.

In order to give a formulation for the CSA-OMHLP taking advantage of these facts, we need to introduce some additional notation. Let I be the number of blocks of consecutive equal non-null elements in $\tilde{\lambda}$ and define the vectors:

1. $\gamma = (\gamma_1, \dots, \gamma_I)$, being $\gamma_i, i = 1, \dots, I$ the value of the elements in the i -th block of repeated elements in $\tilde{\lambda}$.
2. $\alpha = (\alpha_1, \dots, \alpha_I, \alpha_{I+1})$, being α_i with $i = 1, \dots, I$, the number of zero entries between the $(i - 1)$ -th and i -th blocks of positive elements in $\tilde{\lambda}$ and α_{I+1} the number of zeros, if any, after the I -th block of non-null elements in $\tilde{\lambda}$. For notation purposes we define $\alpha_0 = 0$.
3. $\beta = (\beta_1, \dots, \beta_I)$, being $\beta_i, i = 1, \dots, I$ the number of elements in the i -th block of non-null elements in $\tilde{\lambda}$. For the sake of compactness, let $\beta_0 = \beta_{I+1} = 0$.

Next, let denote $\bar{\alpha}_i = \sum_{j=1}^i \alpha_j, \bar{\beta}_i = \sum_{j=1}^i \beta_j$. Moreover, for all $i = 1, \dots, I$ and $h = 1, \dots, G$, let us define the following set of decision variables:

$$u_{ih} = \begin{cases} 1, & \text{if the } (p + \bar{\alpha}_i + \bar{\beta}_{i-1} + 1)\text{-th assignment cost is at least } \hat{c}_{(h)}, \\ 0, & \text{otherwise.} \end{cases}$$

$$v_{ih} = \text{Number of assignments in the } i\text{-th block between the positions } p + \bar{\alpha}_i + \bar{\beta}_{i-1} + 1 \text{ and } p + \bar{\alpha}_i + \bar{\beta}_i \text{ that are at least } \hat{c}_{(h)}.$$

With the above notation the formulation of CSA-OMHLP is:

$$(F_{uv}) \quad \min \quad \sum_{i=1}^I \sum_{h=2}^G \gamma_i (\hat{c}_{(h)} - \hat{c}_{(h-1)}) v_{ih} + \sum_{k \in A} \sum_{\ell \in A} \sum_{m \in A} (\mu c_{k\ell} + \delta c_{\ell m}) s_{k\ell m} \tag{15}$$

s.t. Constraints : (4)–(11),

$$\sum_{i=1}^I \alpha_i u_{ih} + \sum_{i=1}^I v_{ih} + \alpha_{I+1} \geq \sum_{j \in A} \sum_{k \in A} x_{jk}, \quad \forall h = 2, \dots, G \tag{16}$$

$$\hat{c}_{jk} \geq \hat{c}_{(h)}$$

$$u_{ih} \geq u_{i-1,h}, \quad \forall i = 2, \dots, I, h = 1, \dots, G \tag{17}$$

$$\beta_{i-1} u_{ih} \geq v_{i-1,h}, \quad \forall i = 2, \dots, I, h = 1, \dots, G \tag{18}$$

$$v_{ih} \geq \beta_i u_{ih}, \quad \forall i = 1, \dots, I, h = 1, \dots, G \tag{19}$$

$$u_{ih} \in \{0, 1\}, v_{ih} \in \mathbb{Z} \cap [0, \beta_i], \quad \forall i = 1, \dots, I, h = 1, \dots, G \tag{20}$$

$$x_{jk} \in \{0, 1\}, s_{k\ell m} \geq 0, \quad \forall j, k, \ell, m \in A. \tag{21}$$

The objective function (15) is a reformulation of (3) substituting the \bar{u} variables by the new u, v variables and the vector γ , taking advantage of the λ vector properties. Constraints (16) ensure that the number of sites that support a shipping cost to the first hub greater than or equal to $\hat{c}_{(h)}$ is either equal to the number of allocations with a cost at least $\hat{c}_{(h)}$ whenever $v_{ih} > 0$ or less than or equal to α_{I+1} otherwise. Constraints (17) are sorting constraints on the variables u similar to constraints (13), and constraints (18) and (19) provide upper and lower bounds on the variables v depending on the values of variables u .

The main difference between F_u and F_{uv} is that all \bar{u}_{ih} variables associated with blocks of zero λ -values are removed, and those associated with each block of non-null λ values are replaced by $2 \times G$ variables. Therefore, overall we reduce the number of variables by $(n - 2I) \times G$.

Note that in Formulation $F_{\bar{u}}$, the family of constraints that links covering variables (variables \bar{u}) and the allocation variables (variables x), i.e. (12), is given with equalities. This fact implies that the actual dimension of the feasible region in the space of \bar{u} and x variables is smaller than the one that we were currently working on. This is exploited in the new formulation. Indeed, we reduce the number of variables used in the sorting phase replacing \bar{u} by u and v . Therefore, the dimension of the feasible region in the space of u, v, x variables is smaller. In addition, the constraints that link sorting and design variables, namely (16), are given as inequalities. This new representation, although valid for the problem, induces some loss

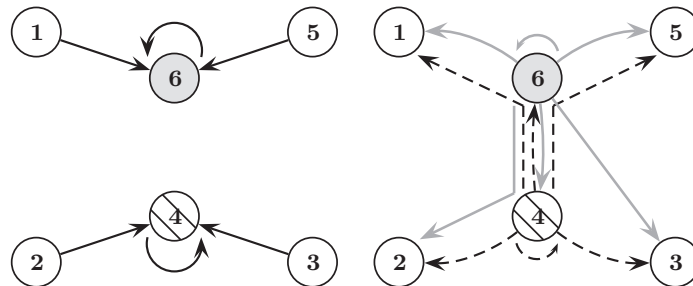


Fig. 1. Illustration of Example 2.1. Left figure represents the allocations of sites to their corresponding first hubs. Right figure represents the flow pattern to the final destinations from the first hubs: 4 (dashed lines) and 6 (grey lines).

of information in that it does not allow us to take full control of the exact number of allocations at some specific cost. This does not affect the resolution process but influences the derivation of valid inequalities.

Finally, for those cases where $\beta_i = 1$ we observe that $v_{ih} = u_{ih}$. This set of constraints whenever valid, was added to reinforce the formulation.

Example 2.1. To illustrate how the F_{uv} versus $F_{\bar{u}}$ formulations work, we consider the following data. Let $A = \{1, \dots, 6\}$ be a set of sites and assume that we are interested in locating $p = 2$ hubs. Let the cost and flow matrices be as follows:

$$C = \begin{pmatrix} 0 & 14 & 15 & 16 & 15 & 9 \\ 5 & 0 & 7 & 2 & 19 & 16 \\ 16 & 5 & 0 & 7 & 1 & 19 \\ 12 & 1 & 10 & 0 & 13 & 1 \\ 1 & 9 & 9 & 15 & 0 & 2 \\ 8 & 10 & 16 & 8 & 4 & 0 \end{pmatrix}, W = \begin{pmatrix} 0 & 15 & 2 & 8 & 11 & 2 \\ 19 & 0 & 1 & 16 & 20 & 7 \\ 3 & 9 & 0 & 3 & 11 & 16 \\ 7 & 2 & 5 & 0 & 14 & 5 \\ 15 & 4 & 20 & 4 & 0 & 1 \\ 12 & 4 & 7 & 11 & 18 & 0 \end{pmatrix}.$$

Therefore, $\hat{c}_{(\cdot)}$, the sorted vector of \hat{c} , is in our case

$\hat{c}_{(\cdot)} = [0, 33, 42, 44, 88, 126, 208, 210, 294, 315, 330, 342, 396, 416, 429, 441, 520, 532, 570, 608, 660, 672, 798, 832, 1008, 1197]$. Hence, $G = 26$. Let $\lambda = (0, 1, 0, 0, 1, 1)$, $\mu = 0.7$, $\delta = 0.9$, and the capacity constraints vector $b = (119, 119, 113, 145, 149, 140)$. The optimal solution opens hubs 4 and 6. The allocation of origin sites to first hub is given by the following values of the variables x (see Fig. 1):

$$x_{16} = x_{24} = x_{34} = x_{44} = x_{56} = x_{66} = 1.$$

Analogously, the allocation of first hubs to final destinations are given by the values of the non null variables s . Thus, the flows considering as first hubs 4 and 6 are (see Fig. 1 for a graphical representation of the delivery paths):

$$s_{442} = 11, s_{443} = 6, s_{444} = 19, s_{461} = 29, s_{465} = 45, s_{466} = 28;$$

$$s_{642} = 23, s_{644} = 23, s_{661} = 27, s_{663} = 29, s_{665} = 29, s_{666} = 3.$$

Moreover, the covering variables \bar{u}_{ih} are given below. Due to their structure, we only report for each i the last one and first zero occurrences since they characterize the remaining values.

$$\begin{aligned} i = 1 &\mapsto \bar{u}_{11} = 1, \bar{u}_{12} = 0 & i = 2 &\mapsto \bar{u}_{21} = 1, \bar{u}_{22} = 0 & i = 3 &\mapsto \bar{u}_{35} = 1, \bar{u}_{36} = 0 \\ i = 4 &\mapsto \bar{u}_{46} = 1, \bar{u}_{47} = 0 & i = 5 &\mapsto \bar{u}_{59} = 1, \bar{u}_{5,10} = 0 & i = 6 &\mapsto \bar{u}_{6,12} = 1, \bar{u}_{6,13} = 0. \end{aligned}$$

The first two assignments are done at a cost $c_{(1)} = 0$, corresponding to the two hubs ($p = 2$). The next assignment has been done at a cost $c_{(5)} = 88$, since $\bar{u}_{35} = 1$ and $\bar{u}_{36} = 0$, and so on. The rest of assignments costs are then $c_{(6)} = 126$, $c_{(9)} = 294$ and $c_{(12)} = 342$.

Hence, the overall cost of this solution is

$$\sum_{i \in A} \sum_{h=2}^G \lambda_i (\hat{c}_{(h)} - \hat{c}_{(h-1)}) \bar{u}_{ih} + \sum_{k \in A} \sum_{\ell \in A} \sum_{m \in A} (\mu c_{k\ell} + \delta c_{\ell m}) s_{k\ell m} = 636 + 1500.8 = 2136.8.$$

In addition, to illustrate how the formulation F_{uv} is related with $F_{\bar{u}}$, we also include the solution of the covering variables u_{ih} and v_{ih} :

Note that we have only one block of repeated non-null elements of the $\tilde{\lambda}$ -vector, so $I = 1$. (See the right part of Fig. 2.) The number of zero entries between two blocks is $\alpha_1 = 2$, and the number of elements in the 1st block of non-null elements is $\beta_1 = 2$. Furthermore, $\gamma_1 = 1$ is the repeated value in the 1st block.

$$\bar{u}_{i,h} = \begin{matrix} & c_{(1)} & c_{(2)} & c_{(3)} & c_{(4)} & c_{(5)} & c_{(6)} & c_{(7)} & c_{(8)} & c_{(9)} & c_{(10)} & c_{(11)} & c_{(12)} & c_{(13)} & \dots & c_{(26)} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & \dots & 0 \end{pmatrix} \end{matrix} \Rightarrow \lambda = \left. \begin{matrix} \left. \begin{matrix} 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{matrix} \right\} \right\} p \\ \left. \left. \begin{matrix} 0 \\ 0 \end{matrix} \right\} \right\} \alpha_1 \\ \left. \left. \left. \begin{matrix} 1 \\ 1 \end{matrix} \right\} \right\} \beta_1 \right\} \xrightarrow{\gamma_1} u_1 \\ v_1 \end{matrix} \right\}$$

Fig. 2. Variables and lambda vector of Example 2.1.

$$\begin{array}{cccccccccccccccccccc}
 & c_{(1)} & c_{(2)} & c_{(3)} & \dots & c_{(9)} & c_{(10)} & c_{(11)} & c_{(12)} & c_{(13)} & c_{(14)} & c_{(15)} & c_{(16)} & c_{(17)} & \dots & c_{(25)} & c_{(26)} \\
 u_{1,h} = & (1 & 1 & 1 & \dots & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0) \\
 v_{1,h} = & (2 & 2 & 2 & \dots & 2 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0)
 \end{array}$$

The variable $u_{1,h}$ points out to the row $p + \alpha_1 + 1 = 5$ of the original variable $\bar{u}_{i,h}$. The variable $v_{1,h}$ accounts for the number of assignments between the positions $p + \alpha_1 + 1 = 5$ and $p + \alpha_1 + \beta_1 = 6$ of $\bar{u}_{i,h}$ that are at least $\hat{c}_{(h)}$. (See Fig. 2.)

From $\bar{u}_{i,h}$, we know that the 5th assignment cost is $\hat{c}_{(9)}$ and the 6th assignment cost is $\hat{c}_{(12)}$. For this reason $v_{1,h}$ is equal to 2 up to column 9, this is the number of assignment costs greater than or equal to $\hat{c}_{(9)}$, this number being equal to 1 from $h = 10$ to $h = 12$, and zero for the remaining columns.

Applying this formulation, the overall reduction in the number of variables is $(n - 2I) \times G = 104$. The rest of variables x and s remain the same, and again the overall cost of this solution is

$$\sum_{i=1}^I \sum_{h=2}^G \gamma_i(\hat{c}_{(h)} - \hat{c}_{(h-1)}) v_{ih} + \sum_{k \in A} \sum_{\ell \in A} \sum_{m \in A} (\mu c_{k\ell} + \delta c_{\ell m}) s_{k\ell m} = 636 + 1500.8 = 2136.8.$$

3. Strengthening the formulation

3.1. Variable fixing

Next, we describe some preprocessing procedures that we have applied to further reduce the size of formulation F_{uv} . We present a number of variable fixing possibilities for the set of variables u and v which are useful in the overall solution process. The variable fixing procedures developed in this section are based on ideas used in Puerto et al. (2011, 2013) but taking advantage of the capacity constraints. Indeed, we are adding the reinforced effective capacity constraints, as well as some surrogated version of constraints (24) since in this aggregated form they give better running times. The preprocessing phase developed in this paper also provides new upper and lower bounds on the variables v . The percentage of variable reduction obtained by these procedures can be found in Table 1 (column named as ‘Fixed’).

Before describing these procedures for fixing variables, the following simple arguments allow us to fix some variables:

1. First, $c_{jj} = 0 \forall j \in A$, i.e., $\hat{c}_{(1)} = 0$. Moreover, every origin where a hub has been located will be allocated to itself as a first hub.
2. Second, $\hat{c}_{jk} \neq 0$ if and only if $j \neq k$, i.e., any non-hub origin is allocated to a first hub at a cost of at least $\hat{c}_{(2)}$.

Therefore, since in this formulation the first p allocations are considered only implicitly, we can fix $u_{i1} = 1$, $v_{i1} = \beta_i$ as well as $u_{i2} = 1$, $v_{i2} = \beta_i$, $\forall i = 1, \dots, I$.

3.1.1. Preprocessing Phase 1: Fixing variables to the upper bounds for the formulation with covering variables strengthen with capacity constraints

Due to the definition of the variables in formulation F_{uv} , one can expect that $u_{ih} = 1$ whenever i is large and h is small to medium size because this would mean that the $(p + \bar{\alpha}_i + \bar{\beta}_{i-1} + 1)$ -th sorted allocation cost would not have been done at cost less than $\hat{c}_{(h)}$. The reader should observe that an analogous strategy applies to the variables v since their interpretation is similar, but in this case the values of v_{ih} would be fixed to β_i . For the cases where it is not possible to fix the corresponding variable v , it could still be possible to establish some lower bounds as we will see later.

Next, to fix variables u_{ih} and v_{ih} for $i = 1, \dots, I$, $h = 1, \dots, G$, we deal with an auxiliary problem that maximizes the number of variables that may assume zero values, satisfying $\hat{c}_{jk} \leq \hat{c}_{(h-1)}$ and the capacity constraints. For any $h = 1, \dots, G$ and $j, k \in A$ such that $\hat{c}_{jk} \leq \hat{c}_{(h-1)}$ let

$$z_{jk}^h = \begin{cases} 1, & \text{if origin site } j \text{ is assigned to hub } k \\ 0, & \text{otherwise.} \end{cases} \tag{22}$$

Table 1
Computational results for formulations F_u vs F_{uv} with preprocessing.

| λ | Formulation F_u | | | Formulation F_{uv} | | | Formulation $F_{uv} + P_1 + P_2$ | | | | | |
|--------------|-------------------|-------|------------------------|----------------------|--------|--------|----------------------------------|--------|----------------------|------------|----------|--------|
| | RGAP | Nodes | Time | RGAP | Nodes | Time | RGAP | Nodes | T_{total} | T_{prep} | Fixed(%) | # Cuts |
| MEDIAN | 13.30 | 234 | 483.9 ⁽¹⁾ | 13.30 | 8139 | 88.4 | 9.90 | 7930 | 85.4 | 6.7 | 4.7 | 509 |
| ANTITRIMMEAN | 17.70 | 8300 | 2523.8 ⁽¹⁸⁾ | 17.60 | 56,000 | 565.8 | 12.10 | 40,522 | 366.3 | 6.7 | 11.8 | 345 |
| TRIMMEAN | 15.00 | 6909 | 1725.3 ⁽⁶⁾ | 13.30 | 24,240 | 184.7 | 12.50 | 15,889 | 117.3 | 6.7 | 2.4 | 163 |
| CENTER | 21.60 | 1501 | 2945.8 ⁽²⁰⁾ | 19.10 | 26,002 | 191.4 | 13.80 | 22,017 | 141.3 | 6.7 | 11.5 | 0 |
| K-CENTRUM | 21.10 | 2534 | 2267.7 ⁽¹⁴⁾ | 16.70 | 28,714 | 312.5 | 11.20 | 20,257 | 200.6 | 6.7 | 3.2 | 359 |
| BLOCKS | 15.80 | 6132 | 1851.6 ⁽⁹⁾ | 15.60 | 61,612 | 1177.7 | 11.60 | 52,644 | 808.1 ⁽⁴⁾ | 6.8 | 19.2 | 297 |
| TOTAL | 17.40 | 4268 | 1966.3 | 15.90 | 34,118 | 420.1 | 11.80 | 26,543 | 286.5 | 6.7 | 8.8 | 279 |

To avoid possible misunderstanding in the cases where variables z_{jk}^h are not defined, i.e. when $\hat{c}_{jk} > \hat{c}_{(h-1)}$, we can assume that $z_{jk}^h := 0$.

For a given h , we introduce the *effective capacity* of a hub k at a cost at most $\hat{c}_{(h-1)}$ as,

$$b_k^{h-1} := \min \left\{ b_k, \sum_{\substack{s \in A \\ \hat{c}_{sk} \leq \hat{c}_{(h-1)}}} W_s \right\}. \tag{23}$$

Indeed, the capacity of a hub k is always lower than or equal to b_k . In addition, if we restrict ourselves to the nodes served at a cost of at most $\hat{c}_{(h-1)}$, then the actual capacity to cover this set should be lower than $\sum_{s \in A: \hat{c}_{sk} \leq \hat{c}_{(h-1)}} W_s$, and this gives us the expression of b_k^{h-1} .

The optimal value $P_1(h)$ of the following problem fixes the maximal number of allocations that may be feasible at a cost of at most $\hat{c}_{(h-1)}$.

$$\begin{aligned} P_1(h) := \max \quad & \sum_{\substack{j,k \in A: \\ \hat{c}_{jk} \leq \hat{c}_{(h-1)}}} z_{jk}^h \\ \text{s.t.} \quad & \sum_{k \in A: \hat{c}_{jk} \leq \hat{c}_{(h-1)}} z_{jk}^h \leq 1, \quad \forall j \in A, \\ & \sum_{j \in A: \hat{c}_{jk} \leq \hat{c}_{(h-1)}} z_{jk}^h \leq ny_k, \quad \forall k \in A \\ & \sum_{k \in A} y_k \leq p \\ & \sum_{j \in A: \hat{c}_{jk} \leq \hat{c}_{(h-1)}} W_j z_{jk}^h \leq b_k^{h-1} y_k, \quad \forall k \in A \\ & z_{jk}^h, y_k \in \{0, 1\}, \quad \forall j, k \in A, \forall h = 1, \dots, G. \end{aligned} \tag{24}$$

Depending on the value $P_1(h)$ we can fix some variables to their upper bounds. Let us denote by $i_1(h) \in \{1, \dots, I\}$ the index such that

$$p + \bar{\alpha}_{i_1(h)-1} + \bar{\beta}_{i_1(h)-1} < P_1(h) \leq p + \bar{\alpha}_{i_1(h)} + \bar{\beta}_{i_1(h)}.$$

Then,

$$\left\{ \begin{array}{l} u_{ih} = 1, v_{ih} = \beta_i, i = i_1(h), \dots, I, \\ \left\{ \begin{array}{l} v_{i_1(h),h} \geq p + \bar{\alpha}_{i_1(h)} + \bar{\beta}_{i_1(h)} - P_1(h), \\ u_{ih} = 1, v_{ih} = \beta_i, i = i_1(h) + 1, \dots, I, \end{array} \right. \end{array} \right\} \text{ otherwise.}$$

3.1.2. Preprocessing Phase 2: Fixing variables to their lower bounds

Following similar argument to the previous subsection, one can expect that many variables u and v in the top-right hand corner of the matrices of variables u and v , respectively, will take value 0 in the optimal solution. Indeed, $u_{ih} = 0$ means that the $(p + \bar{\alpha}_i + \bar{\beta}_{i-1} + 1)$ -th sorted allocation cost is less than $\hat{c}_{(h)}$ which is very likely to be true if h is sufficiently large and i is small. Note that an analogous strategy applies to the variables v since their interpretation is similar. For the cases where it is not possible to fix the corresponding variable v it could be still possible to establish some upper bounds.

For any $h = 2, \dots, G, j, k \in A$ such that $\hat{c}_{jk} \geq \hat{c}_{(h-1)}$ let define variables z_{jk}^h as (22). To avoid possible misunderstanding in the cases where variables z_{jk}^h are not defined, i.e. when $\hat{c}_{jk} < \hat{c}_{(h-1)}$, we can assume that $z_{jk}^h := 0$. Using these variables, the formulation of the problem that maximizes the number of non-fixed allocations at a cost at most $\hat{c}_{(h-1)}$ is:

$$\begin{aligned} P_2(h) := \max \quad & \sum_{\substack{j,k \in A: \\ \hat{c}_{jk} \geq \hat{c}_{(h-1)}}} z_{jk}^h \\ \text{s.t.} \quad & \sum_{k \in A: \hat{c}_{jk} \geq \hat{c}_{(h-1)}} z_{jk}^h \leq 1, \quad \forall j \in A \\ & \sum_{j \in A: \hat{c}_{jk} \geq \hat{c}_{(h-1)}} z_{jk}^h \leq ny_k, \quad \forall j, k \in A \\ & \sum_{k \in A} y_k \leq p, \\ & z_{jk}^h, y_k \in \{0, 1\}, \quad \forall j, k \in A, \forall h = 1, \dots, G. \end{aligned} \tag{25}$$

Note that the value $P_2(h)$ implies that there are no feasible solutions of the original problem with less than $n - P_2(h)$ allocations fixed at a cost at most $\hat{c}_{(h)}$.

Let $1 \leq i_2(h) \leq I$ be the index such that

$$p + \bar{\alpha}_{i_2(h)-1} + \bar{\beta}_{i_2(h)-1} < n - P_2(h) \leq p + \bar{\alpha}_{i_2(h)} + \bar{\beta}_{i_2(h)}.$$

Thus, in any feasible solution of the problem we have that:

$$\left\{ \begin{array}{l} u_{ih} = 0, v_{ih} = 0, \quad i = 1, \dots, i_2(h) - 1, \\ \left\{ \begin{array}{l} u_{i_2(h),h} = 0, \quad v_{i_2(h),h} \leq p + \bar{\alpha}_{i_2(h)} + \bar{\beta}_{i_2(h)} - (n - P_2(h)), \\ u_{ih} = 0, \quad v_{ih} = 0, \quad i = 1, \dots, i_2(h) - 1, \end{array} \right. \end{array} \right\} \quad \begin{array}{l} \text{if } n - P_2(h) \leq p + \bar{\alpha}_{i_2(h)} + \bar{\beta}_{i_2(h)-1} \\ \text{otherwise.} \end{array}$$

Note that, whenever $n - P_2(h) = p$, there is nothing to fix and therefore no variables are set to zero in column h .

3.2. Valid inequalities

In order to strengthen formulation F_{uv} we have studied several families of valid inequalities. In fact, taking advantage of previous experience on the non-capacitated version of the problem we have borrowed a first family of valid inequalities that are very simple and that have proven to be effective in different ordered median problems with covering variables (Puerto et al., 2011, 2013). This family is

$$u_{ih} \geq u_{i,h+1}, \quad i = 1, \dots, I, \quad h = 1, \dots, G - 1, \tag{26}$$

$$v_{ih} \geq v_{i,h+1}, \quad i = 1, \dots, I, \quad h = 1, \dots, G - 1. \tag{27}$$

Since these families are straightforward consequence of the definition of variables u and v , they have been included in the original formulation.

In the following, we describe several alternative families of valid inequalities: three sets of inequalities, (28), (29)–(32), and (33)–(36), based on the combination of ordering and capacity requirements and two more sets, (37)–(39), that do not use capacities.

3.2.1. First family of valid inequalities: valid inequalities based on capacity I

We can add several families of valid inequalities based on capacity issues to this model that help in solving the problem by reducing the gap of the linear relaxation and the CPU time to explore the branch and bound search tree.

Observe that, the capacity of the set of hubs that may be used to assign origins in A at a cost at most $\hat{c}_{(h-1)}$, is given by

$$\sum_{k \in A} b_k^{h-1} x_{kk}$$

where b_k^{h-1} is the effective capacity at a cost at most $\hat{c}_{(h-1)}$, defined by (23). Recall that, although the capacity of a hub k is always lower than or equal to b_k , when we restrict to the nodes served at a cost of at most $\hat{c}_{(h-1)}$, then the actual capacity to cover this set should be lower than b_k^{h-1} . Making use of the above observation, we can add the following family of constraints as valid inequalities

$$\sum_{\substack{j \in A \\ \hat{c}_{jk} \leq \hat{c}_{(h-1)}}} W_j x_{jk} \leq b_k^{h-1} x_{kk} \quad \forall h = 2, \dots, G, k \in A \tag{28}$$

which enforces that all the flow sent from origin-hubs at a cost at most $\hat{c}_{(h-1)}$ cannot exceed the effective capacity at that cost. Observe that in the case where b_k^{h-1} takes the value b_k , (11) dominates (28), but in the case where $b_k^{h-1} = \sum_{s \in A: \hat{c}_{sk} \leq \hat{c}_{(h-1)}} W_s$, (28) becomes $\sum_{j \in A: \hat{c}_{jk} \leq \hat{c}_{(h-1)}} W_j x_{jk} \leq \sum_{j \in A: \hat{c}_{jk} \leq \hat{c}_{(h-1)}} W_j x_{kk}$. Observe that this last valid inequality is an alternative surrogation, with capacity coefficients, of constraints $x_{jk} \leq x_{kk}$ that although valid do not appear in the model because of its large cardinality. This new form of aggregation has provided good results in the computational experiments.

3.2.2. Second family of valid inequalities: valid inequalities based on capacity II

This section introduces another family of valid inequalities based on capacity issues that help in solving the problem. In order to present these new valid inequalities based on capacity requirements we introduce some new notation. Assume without loss of generality that $W_i \leq W_{i+1}$ for $i \in A \setminus \{n\}$ and let $\bar{W}_j = \sum_{r=1}^j W_r$ and $S_k := \{i \in A : i \leq k\}$ for $k \in A$ be a given set of origin sites.

In case that the effective capacity at a cost at most $\hat{c}_{(h-1)}$ is not sufficient to cover the demand of S_k , i.e. $\sum_{j=1}^n (b_j^{h-1} - W_j) x_{jj}$ is less than $\sum_{s=1}^k W_s (1 - x_{ss})$, then at most $k - 1$ origins of S_k can be allocated at a cost lower than or equal to $\hat{c}_{(h-1)}$. This argument can be applied for each h to the corresponding $\hat{c}_{(h-1)}$ value. Moreover, we have chosen this particular structure of S_k consisting of the k origins (nodes) with the k -smallest flows, because given a fixed amount of flow, the maximal cardinality

set of origins such that the overall flow originated in this set is lower than or equal to this amount is provided by a set S_k for some $k \in A$. Therefore, since we are dealing with the worst cases, it allows us to fix some variables u and v through the following valid inequalities. For each $h = 2, \dots, G$ we obtain the following:

- If $\bar{\alpha}_{i-1} + \bar{\beta}_{i-1} < k \leq \bar{\alpha}_i + \bar{\beta}_{i-1} + 1$, for some $i \in \{1, \dots, I\}$, namely if the index of the last element, k , that defines S_k lies in the i -th block of null elements in the $\tilde{\lambda}$ vector, then

$$\bar{W}_k u_{ih} + \sum_{j \in A} (b_j^{h-1} - W_j) x_{ij} \geq \sum_{s=1}^k W_s (1 - x_{ss}). \tag{29}$$

Observe that the above inequality amounts to a disjunctive condition: either the effective capacity at a cost at most $\hat{c}_{(h-1)}$ is enough to cover the demand of the k smallest flows from origin sites or the i -th sorted cost allocation must be assigned at a cost at least $c_{(h)}$.

- If $\bar{\alpha}_i + \bar{\beta}_{i-1} + 1 < k \leq \bar{\alpha}_i + \bar{\beta}_i$, and namely if the index of the last element, k , that defines S_k lies in the i -th block of non-null elements, then

$$\bar{W}_k v_{ih} + \sum_{j \in A} (b_j^{h-1} - W_j) x_{ij} \geq \sum_{s=1}^k W_s (1 - x_{ss}). \tag{30}$$

In this case, the inequality is similar to the previous one but written in terms of the variables v that allow to control the capacity whenever k falls within a block of non-null elements in the $\tilde{\lambda}$ vector.

Remark 3.1. Recall that if $k > P_1(h) - p$, variables u_{ih} and v_{ih} have been already fixed by the Preprocessing Phase 1, and for the above inequality to be effective $k \leq P_1(h) - p$, or equivalently, $i \leq i_1(h)$.

Based on the same arguments we can add a larger family of valid inequalities built on arbitrary sets of origin sites. Let S be a set of origin sites, and suppose that $A_S = \sum_{s \in S} W_s$ satisfies $\bar{W}_k \leq A_S < \bar{W}_{k+1}$,

- If $\bar{\alpha}_{i-1} + \bar{\beta}_{i-1} < k \leq \bar{\alpha}_i + \bar{\beta}_{i-1} + 1$, for some $i \in \{1, \dots, I\}$, and $k \leq P_1(h) - p$ then

$$A_S u_{ih} + \sum_{j \in A} (b_j^{h-1} - W_j) x_{ij} \geq \sum_{s=1}^k W_s (1 - x_{ss}). \tag{31}$$

- If $\bar{\alpha}_i + \bar{\beta}_{i-1} + 1 < k \leq \bar{\alpha}_i + \bar{\beta}_i$, and $k \leq P_1(h) - p$ then

$$A_S v_{ih} + \sum_{j \in A} (b_j^{h-1} - W_j) x_{ij} \geq \sum_{s=1}^k W_s (1 - x_{ss}). \tag{32}$$

Now, assuming a more general case and for an improvement of the above inequalities (29) and (30), for a given $h \in \{2, \dots, G\}$, and $k \leq P_1(h) - p$, $k \in \{1, \dots, \bar{\alpha}_{i+1} + \bar{\beta}_i\}$, let

$$M_s := \min_{j(\neq s) \in A} \hat{c}_{js}.$$

M_s is the minimum allocation cost to s as an open hub. In other words, no allocation to hub s is possible at a cost less than M_s , except in case s were a hub itself. We shall call this value the *empty radius of s* .

Define $s(h-1, k)$ to be the index of the sorted sequence of elements W_s such that there are exactly k elements W_s with $s \leq s(h-1, k)$ and $M_s \leq \hat{c}_{(h-1)}$, namely $s(h-1, k)$, is the index such that

$$|\{s : s \leq s(h-1, k); M_s \leq \hat{c}_{(h-1)}\}| = k.$$

Then it holds that,

- If $\bar{\alpha}_{i-1} + \bar{\beta}_{i-1} < k \leq \bar{\alpha}_i + \bar{\beta}_{i-1} + 1$, for some $i \in \{1, \dots, I\}$, and $k \leq P_1(h) - p$

$$\left(\bar{W}_{s(h-1, k)} - \sum_{\substack{s=1 \\ M_s > \hat{c}_{(h-1)}}}^{s(h-1, k)} W_s \right) u_{ih} + \sum_{j \in A} (b_j^{h-1} - W_j) x_{ij} \geq \sum_{\substack{s=1 \\ M_s \leq \hat{c}_{(h-1)}}}^{s(h-1, k)} W_s (1 - x_{ss}) - \sum_{\substack{s=1 \\ M_s > \hat{c}_{(h-1)}}}^{s(h-1, k)} (W_{s(h-1, k)} - W_s) x_{ss}. \tag{33}$$

The above inequality is also a disjunctive condition that reinforces the family of valid inequalities (29). It states that if the effective capacity at a cost at most $\hat{c}_{(h-1)}$ is not enough to cover the flow sent from origin sites that are not hubs and that can be allocated at some costs less than or equal to $\hat{c}_{(h-1)}$ then some of the origin sites with allocation costs less than $\hat{c}_{(h-1)}$

must be assigned at a cost at least $\hat{c}_{(h)}$. We observe that the use of u variables in the left-hand side of the inequality is due to the fact that k falls within a block of null elements in the $\tilde{\lambda}$ vector. A similar inequality also holds when k falls within a block of non-null elements as shown below.

- If $\bar{\alpha}_i + \bar{\beta}_{i-1} + 1 < k \leq \bar{\alpha}_i + \bar{\beta}_i$, and $k \leq P_1(h) - p$ then

$$\left(\bar{W}_{s(h-1,k)} - \sum_{\substack{s=1 \\ M_s > \hat{c}_{(h-1)}}}^{s(h-1,k)} W_s \right) v_{ih} + \sum_{j \in A} (b_j^{h-1} - W_j) x_{jj} \geq \sum_{\substack{s=1 \\ M_s \leq \hat{c}_{(h-1)}}}^{s(h-1,k)} W_s (1 - x_{ss}) - \sum_{\substack{s=1 \\ M_s > \hat{c}_{(h-1)}}}^{s(h-1,k)} (W_{s(h-1,k)} - W_s) x_{ss}. \tag{34}$$

This inequality is similar to the previous one whenever the index k falls within a block of non-null elements in the $\tilde{\lambda}$ vector.

Finally, as the index $s(h-1, k)$ should be greater than or equal to k , we can split the above equations, (33) and (34), into $s(h-1, k) - k$ equivalent inequalities. This is, for any $t = k, \dots, s(h-1, k)$, define $\hat{s}(h-1, k, t)$ to be the index of the sorted sequence of elements W_s such that

$$|\{s; s \leq \hat{s}(h-1, k, t); M_s \leq \hat{c}_{(h-1)}\}| + t - \hat{s}(h-1, k, t) = k.$$

Then it holds that,

- If $\bar{\alpha}_{i-1} + \bar{\beta}_{i-1} < k \leq \bar{\alpha}_i + \bar{\beta}_{i-1} + 1$, for some $i \in \{1, \dots, I\}$, and $k \leq P_1(h) - p$

$$\left(\bar{W}_t - \sum_{\substack{s=1 \\ M_s > \hat{c}_{(h-1)} \\ \text{and } s \leq \hat{s}(h-1,k,t)}}^t W_s \right) u_{ih} + \sum_{j \in A} (b_j^{h-1} - W_j) x_{jj} \geq \sum_{\substack{s=1 \\ M_s \leq \hat{c}_{(h-1)} \\ \text{or } s > \hat{s}(h-1,k,t)}}^t W_s (1 - x_{ss}) - \sum_{\substack{s=1 \\ M_s > \hat{c}_{(h-1)} \\ \text{and } s \leq \hat{s}(h-1,k,t)}}^t (W_t - W_s) x_{ss}. \tag{35}$$

- If $\bar{\alpha}_i + \bar{\beta}_{i-1} + 1 < k \leq \bar{\alpha}_i + \bar{\beta}_i$ and $k \leq P_1(h) - p$, then

$$\left(\bar{W}_t - \sum_{\substack{s=1 \\ M_s > \hat{c}_{(h-1)} \\ \text{or } s \leq \hat{s}(h-1,k,t)}}^t W_s \right) v_{ih} + \sum_{j \in A} (b_j^{h-1} - W_j) x_{jj} \geq \sum_{\substack{s=1 \\ M_s \leq \hat{c}_{(h-1)} \\ \text{or } s > \hat{s}(h-1,k,t)}}^t W_s (1 - x_{ss}) - \sum_{\substack{s=1 \\ M_s > \hat{c}_{(h-1)} \\ \text{or } s \leq \hat{s}(h-1,k,t)}}^t (W_t - W_s) x_{ss}. \tag{36}$$

Observe that if $t = s(h-1, k)$ then $\hat{s}(h-1, k, t) = s(h-1, k)$. Thus, the families of valid inequalities (35) and (36) include as particular instances the families (33) and (34).

3.2.3. Third family of valid inequalities: disjunctive implications

The third family of valid inequalities, directly borrowed from Puerto et al. (2013), state disjunctive implications on the origin-first hub allocation costs. They ensure that either origin site j is allocated to a first hub at a cost of at least $\hat{c}_{(h)}$ or there is an open hub k such that $\hat{c}_{jk} < \hat{c}_{(h)}$. This argument can be formulated through the following family of valid inequalities:

$$\sum_{k \in A; \hat{c}_{jk} \geq \hat{c}_{(h)}} x_{jk} + \sum_{k \in A; \hat{c}_{jk} < \hat{c}_{(h)}} x_{kk} \geq 1, \quad \forall j \in A, h = 1, \dots, G. \tag{37}$$

3.2.4. Fourth family of valid inequalities

Using the definition of the variables u and v , we establish a lower and an upper bound of the number of feasible allocations at a cost $\hat{c}_{(h-1)}$. Observe that using the family of constraints (12) for the original formulation, the exact number of allocations done at a cost $\hat{c}_{(h-1)}$ is given by $\sum_{i \in A} (\bar{u}_{i,h-1} - \bar{u}_{i,h})$. However, since in formulation F_{uv} the number of variables has been considerably reduced, some information is lost. In particular, we cannot keep under control with this new formulation the exact number of allocations at a cost $\hat{c}_{(h-1)}$. Indeed, the counterpart to equalities (12) in formulation F_{uv} is the family of constraints (16). Therefore, we are only able to give a lower and upper bound on this number of allocations. These lower and upper bounds are formulated, respectively, by the following two families of constraints:

$$\sum_{j \in A} \sum_{\substack{k \in A \\ \hat{c}_{jk} = \hat{c}_{(h-1)}}} x_{jk} \geq \sum_{i=1}^I (v_{i,h-1} - v_{i,h}) + \sum_{i=2}^I \alpha_i (u_{i-1,h-1} - u_{i,h}), \quad \forall h = 2, \dots, G, \tag{38}$$

$$\sum_{j \in A} \sum_{\substack{k \in A \\ \hat{c}_{jk} = \hat{c}_{(h-1)}}} x_{jk} \leq \sum_{i=1}^I (v_{i,h-1} - v_{i,h}) + \sum_{i=1}^I \alpha_i (u_{i,h-1} - u_{i,h}) + (1 - u_{Ih}) \alpha_{I+1} + \alpha_1 u_{1h} + \sum_{i=1}^{I-1} \alpha_{i+1} (u_{i+1,h} - u_{i,h}), \quad \forall h = 2, \dots, G. \tag{39}$$

The first sum in the right hand side of both families gives the exact number of allocations at cost $\hat{c}_{(h-1)}$ in the positions corresponding to non-null blocks of the vector λ . However, the second sum in the right hand side of constraints (38) provides a lower bound of the number of allocations at cost $\hat{c}_{(h-1)}$ in the positions corresponding to the null blocks. In the same way, the 2nd to the 4th sums in (39) provide an upper bound on the number of these allocations.

4. Computational results

The formulations given to the CSA-OMHLP with the corresponding strengthening and preprocessing phases, described in this paper, were implemented in the commercial solver XPRESS IVE 1.23.02.64 running on a Intel(R) Core(TM) i5-3450 CPU @3.10 GHz 6 GB RAM.

The cut generation option of XPRESS was disabled in order to compare the relative performance of the formulations cleanly.

For this purpose we use the AP data set publicly available at <http://www.cmis.csiro.au/or/hubLocation> (see Ernst and Krishnamoorthy, 1996). As in previous papers on the field related to the uncapacitated version of this problem, we tested the formulations on a testbed of five instances for each combination of costs matrices varying: (i) n in $\{15, 20, 25, 28, 30\}$, (ii) three different values of p ($p \in \{3, 5, 8\}$ for $n = 15$ and $p \in \{3, 8, 10\}$ for $n \in \{20, 25, 28, 30\}$) and (iii) $\mu = 0.7$, $\delta = 0.9\mu$ and six different λ -vectors. These λ -vectors are the well-known Median $\lambda = (1, \dots, 1)$, Anti- $(k_1 + k_2)$ -trimmed-mean $\lambda = (1, \dots, 1, 0, \dots, 0, 1, \dots, 1, 0, \dots, 0)$, $(k_1 + k_2)$ -Trimmed-mean $\lambda = (0, \dots, 0, 1, \dots, 1, 0, \dots, 0)$, with $k_1 = k_2 = \lceil 0.2n \rceil$, Center $\lambda = (0, \dots, 0, 1)$, and k -Centrum $\lambda = (0, \dots, 0, 1, \dots, 1)$ with $k = \lceil 0.2n \rceil$, as well as a $\{0, 1\}$ -blocks λ -vector (three alternate $\{0 - 1\}$ -blocks of lambda weights, i.e. $\lambda = (0, \dots, 0, 1, \dots, 1, 0, \dots, 0, 1, \dots, 1, 0, \dots, 0, 1, \dots, 1)$). Therefore, for the each combination of n , p and λ we have tested five instances. This is, a total number of 450 instances have been used to test the performance of the proposed models.

The capacities were randomly generated in $[\min_i W_i, (1/2)\sum_{i \in A} W_i]$. This generation procedure does not ensure in all cases feasible instances, as capacity constraints can be very tight for problems with a low number of hubs ($n = 3, 5$). Overall, in our experiments we got, initially 10 infeasible instances out of 75 (13.3%). These instances were replaced by new feasible ones (generated with the same capacity structure). Hence, the reader may observe that the generation procedure gives tight capacity constraints.

First of all, and for the sake of readability, we present in Fig. 3 a summary of our computational results. A more detailed report of those results is also included in Tables 1–3 and a full description of these tables can be seen in <http://arxiv.org/abs/1503.04468>.

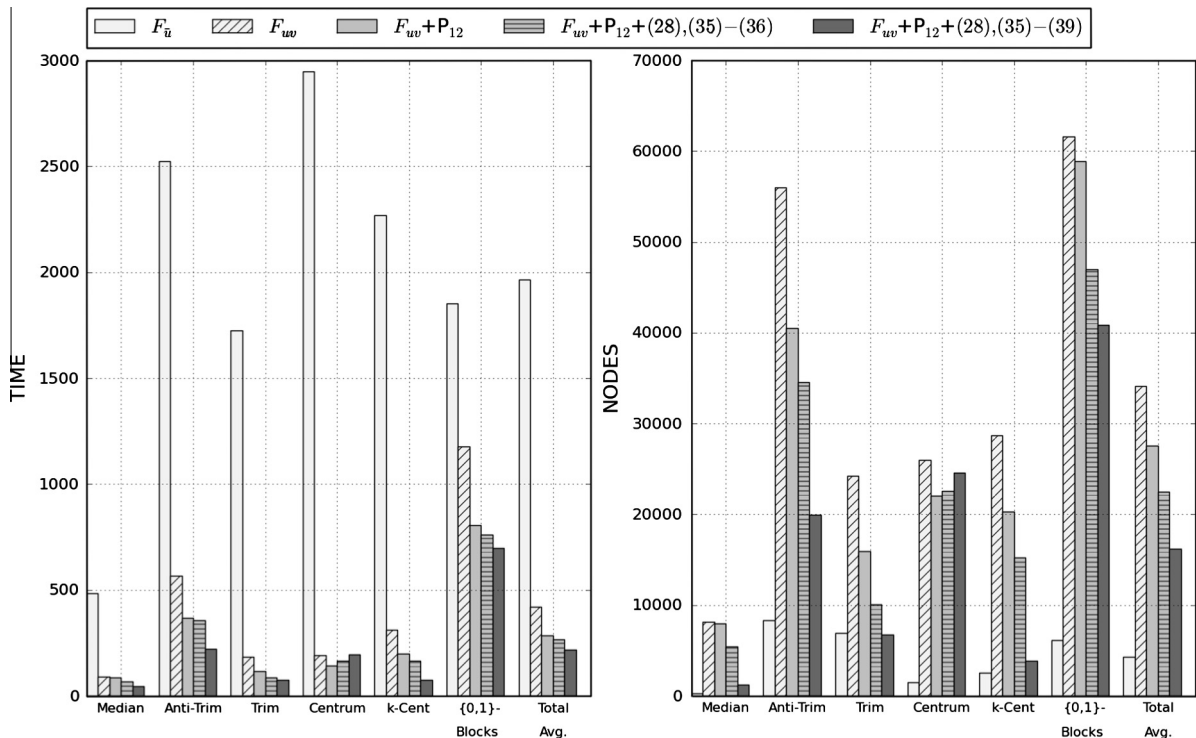


Fig. 3. Summary of computational results carried out in the paper.

Table 2
Computational results for Formulation F_{uv} with Valid Inequalities (I).

| λ | $F_{uv} + P_1 + P_2 = P_{12}$ | | | $P_{12} + (28)-(30)$ | | | $P_{12} + (28), (33), (34)$ | | | $P_{12} + (28), (35), (36)$ | | |
|---------------|-------------------------------|--------|----------------------|----------------------|--------|----------------------|-----------------------------|--------|----------------------|-----------------------------|--------|----------------------|
| | RGAP | Nodes | time | RGAP | Nodes | time | RGAP | Nodes | time | RGAP | Nodes | time |
| MEDIAN | 9.90 | 7930 | 85.4 | 9.40 | 5867 | 73.9 | 9.40 | 5251 | 66.7 | 9.40 | 5450 | 66.8 |
| ANTI-TRIMMEAN | 12.10 | 40,522 | 366.3 | 11.90 | 29,007 | 279.3 | 11.90 | 29,298 | 291.8 | 11.90 | 34,543 | 355.4 |
| TRIMMEAN | 12.50 | 15,889 | 117.3 | 12.30 | 9706 | 84.1 | 12.30 | 9846 | 84.1 | 12.30 | 10,062 | 85.5 |
| CENTER | 13.8 | 22,017 | 141.3 | 13.8 | 22,542 | 152.7 | 13.80 | 24,424 | 167.9 | 13.80 | 22,594 | 164.6 |
| K-CENTRUM | 11.20 | 20,257 | 200.6 | 11.10 | 18,888 | 199.0 | 11.10 | 22,669 | 256.9 | 11.10 | 15,194 | 164.4 |
| BLOCKS | 11.60 | 52,644 | 808.1 ⁽⁴⁾ | 11.50 | 52,303 | 858.9 ⁽⁵⁾ | 11.50 | 46,687 | 757.2 ⁽⁶⁾ | 11.50 | 46,981 | 760.8 ⁽⁴⁾ |
| TOTAL | 11.80 | 26,543 | 286.5 | 11.70 | 23,052 | 274.7 | 11.70 | 23,029 | 270.8 | 11.70 | 22,470 | 266.2 |

Table 3
Computational results for Formulation F_{uv} with Valid Inequalities (II).

| λ | $P_{12} + (28), (35), (36)$ | | | $P_{12} + (28), (35)-(37)$ | | | $P_{12} + (28), (35)-(38)$ | | | $P_{12} + (28), (35)-(39)$ | | |
|---------------|-----------------------------|--------|----------------------|----------------------------|--------|----------------------|----------------------------|--------|----------------------|----------------------------|--------|----------------------|
| | RGAP | Nodes | time | RGAP | Nodes | time | RGAP | Nodes | time | RGAP | Nodes | time |
| MEDIAN | 9.40 | 5450 | 66.8 | 6.10 | 1263 | 41.9 | 6.10 | 1203 | 40.9 | 6.10 | 1231 | 43.2 |
| ANTI-TRIMMEAN | 11.90 | 34,543 | 355.4 | 10.90 | 19,869 | 229.1 | 10.90 | 18,915 | 223.1 | 10.90 | 19,935 | 222.2 |
| TRIMMEAN | 12.30 | 10,062 | 85.5 | 11.50 | 7364 | 75.0 | 11.50 | 6699 | 73.1 | 11.50 | 6754 | 73.3 |
| CENTER | 13.80 | 22,594 | 164.6 | 13.60 | 23,200 | 176.1 | 13.60 | 22,082 | 178.4 | 13.60 | 24,560 | 194.5 |
| K-CENTRUM | 11.10 | 15,194 | 164.4 | 9.70 | 5618 | 83.3 | 9.50 | 9294 | 147.6 | 9.60 | 3875 | 75.1 |
| BLOCKS | 11.50 | 46,981 | 760.8 ⁽⁴⁾ | 10.30 | 39,332 | 784.0 ⁽⁴⁾ | 10.30 | 41,019 | 720.7 ⁽³⁾ | 10.30 | 40,829 | 697.4 ⁽³⁾ |
| TOTAL | 11.70 | 22,470 | 266.2 | 10.40 | 16,108 | 231.6 | 10.30 | 16,536 | 230.7 | 10.30 | 16,197 | 217.6 |

Fig. 3 shows average results for each one of the considered problem types (different λ -vectors). The left chart refers to average CPU times and the right chart to the number of explored nodes of the B&B tree. Both charts contain the same number of blocks standing for the different types of λ -vectors plus and additional block (Total Avg.) for the consolidated average of all λ -vectors. Each block compares the behavior of the different formulations ($F_{\bar{u}}$, F_{uv}) and their strengthening (variable fixing and valid inequalities). The heading explains the meaning of bars within each block: $F_{\bar{u}}$ and F_{uv} stand for the corresponding formulations; $F_{uv} + P_{12}$ when the two preprocesses $P_1(h)$ and $P_2(h)$ are applied; and $F_{uv} + P_{12} + (VI)$ when the Valid Inequalities are added as well, denoted with their corresponding references.

Analyzing this figure we observe the improvement obtained with F_{uv} and its strengthening as compared with $F_{\bar{u}}$ or even F_{uv} alone. Actually, the overall reduction in running time with respect to the initial formulation is around 89%.

Next, focusing in the best model, namely $F_{uv} + P_{12} + (28), (35)-(39)$, we observe that the most time consuming problem is the one with λ -vector given by $\{0, 1\}$ -blocks with a significant difference with respect to the remaining λ -vectors. The second most time consuming problem corresponds with the Antitrimmean. Similar conclusions, regarding the number of nodes, are obtained looking at the right chart of Fig. 3. It is worth mentioning that although the formulation $F_{\bar{u}}$ provides the worst computational times, it is the one that reports the lowest number of nodes in the B&B tree. This fact is explained because $F_{\bar{u}}$ has a larger number of variables than F_{uv} which results in more difficult LP relaxations in each node of the B&B tree.

In spite of that Fig. 3 shows the general overview of our computational results, we also include in the following a more detailed analysis based on Tables 1–3 (a full version of the paper including all the tables with the computational results for each value of n and p can be seen in <http://arxiv.org/abs/1503.04468>).

Table 1 reports the results of the formulations and different preprocessing phases developed in this paper for the CSA-OMHLP. The first three columns correspond to some of the computational results obtained by solving the CSA-OMHLP with $F_{\bar{u}}$ formulation. The next three columns correspond to Formulation F_{uv} , and the rest of the columns to the latter formulation plus the two preprocessing procedures, i.e. $F_{uv} + P_1 + P_2$. Columns *RGAP*, *Nodes* and *Time* stand for the averages of: the gap in the root node, number of nodes in the B&B tree and the CPU time in seconds; the time was limited to two hours of CPU. To obtain a general idea of the comparisons among these averaged values, for the results in the column *Time* and for different formulations and/or valid inequalities applied, we have accounted the value 7200 s for those instances that exceed the time limit. A superindex in their corresponding averaged time value states the number of instances exceeding the CPU time limit; in the same way, the values used to computed the average of the column *Nodes* have been the number of nodes of the B&B tree when the CPU time limit was reached.

The column *Fixed* gives the percentage of variables that have been fixed after the Preprocessing Phases 1 and 2. Column *Cuts* provides the number of the lower and upper bounds over the variables v added to the model, after running the corresponding preprocessing phases. Finally, T_{prep} reports the CPU time in seconds of the corresponding preprocessing phases and column T_{total} reports the overall CPU time in seconds to solve the problem including the corresponding preprocessing phase. The row *TOTAL* provides the averaged results among all the tested instances.

Table 1 shows, that 68 instances required more than two hours to be solved using Formulation $F_{\bar{u}}$, however the remaining two analyzed ways to solve this problem were able to solve all the instances within the CPU time limit, except in a few cases

Table 4
Summary of the improvements presented in this paper.

| λ | Imp_ $F_{\bar{u}}-F_{uv}$ (%) | Imp_ F_{uv} (%) | Imp_TOTAL (%) |
|---------------|-------------------------------|-------------------|---------------|
| MEDIAN | 81.73 | 51.13 | 91.07 |
| ANTI-TRIMMEAN | 77.58 | 60.73 | 91.20 |
| TRIMMEAN | 89.29 | 60.31 | 95.75 |
| CENTER | 93.50 | 26.18 | 95.20 |
| K-CENTRUM | 86.22 | 75.97 | 96.69 |
| BLOCKS | 36.40 | 40.78 | 62.34 |
| TOTAL | 77.45 | 52.52 | 88.71 |

for the $\{0, 1\}$ -blocks problem types. Regarding the running times for the different types of problems, we can see that Anti-TrimMean and Blocks have been the problems that need more time to be solved (for the new formulation F_{uv}). In any case, we observed that there is a considerable reduction of the running times from the original formulation to the improved formulation, and after applying the preprocessing phases. From Table 1, one can remark that Formulation F_{uv} with Preprocessing Phases 1 and 2 provides better results, with a reduction of around a 86% of the time with respect to Formulation $F_{\bar{u}}$ (taking into account that the latter formulation was not able to solve all the studied instances before the time limit was exceeded). Moreover, the reduction of the running times with respect to Formulation F_{uv} (without preprocessing phases) is around 32%.

As for the comparison between RGAP's, we observe that the average gap of the linear relaxation after preprocessing reduces around 32% from the original formulation (Formulation $F_{\bar{u}}$) to Formulation F_{uv} with Preprocessing Phases 1 and 2. In any case, it is also worth noting that the gap from the original formulation to the improved formulation (without any preprocessing phase) is reduced around 9%, what implies that even though this improved formulation uses much less number of variables, it provides a better RGAP.

Table 2 presents several improvements to Formulation F_{uv} with Preprocessing Phases 1 and 2. In particular, the second, third and fourth blocks of columns summarize the results when the combination of valid inequalities (28)–(30); (28), (33), (34) and (28), (35), (36) are added to this formulation, respectively. In general, we can see that the latter combination provides the best results. In particular, there is an improvement of 7% of the running times with respect to Formulation F_{uv} with Preprocessing Phases 1 and 2. Moreover, with respect to the number of nodes the improvement is around 15%, but the RGAP is similar for all the analyzed reinforcements.

Since the best behavior observed was obtained with reinforcement given by (28), (35), (36), in the rest of our tests we have used this configuration to make further strengthening. Table 3 presents several improvements to the Formulation F_{uv} with Preprocessing Phases 1 and 2 and valid inequalities (28), (35), (36). In particular, the second, third and fourth blocks of columns report the results after adding to the latter configuration the family of valid inequalities (37), (37), (38) and (37)–(39), respectively. This table shows that the best results are obtained when all valid inequalities, (37)–(39), are added to the current configuration; with an improvement of 18% in the running time, 28% in the number of nodes and 12% in the RGAP.

The overall conclusion of our experiment is that in order to solve CSA-OMHLP the best combination of formulation and strengthening is to use $F_{uv} + P_{12} + (28) + (35)–(39)$. This configuration allows to solve medium size instances within 10 min of CPU time. In summary, we have obtained: (i) an average improvement of 77.45% of CPU time by using formulation F_{uv} instead of $F_{\bar{u}}$ (see column Imp_ $F_{\bar{u}}-F_{uv}$ of Table 4); (ii) an average improvement of 52.52% of CPU time over the formulation F_{uv} by using the preprocessing phases for fixing variables and the valid inequalities presented in this paper (see column Imp_ F_{uv} of Table 4), and (iii) an overall improvement, on average, of 88.71% from the initial formulation $F_{\bar{u}}$ to $F_{uv} + P_{12} + (28) + (35)–(39)$, see column Imp_TOTAL of Table 4.

5. Concluding remarks

This paper can be considered as an initial attempt to address the capacitated single-allocation ordered hub location problem. The formulations, strengthening and preprocessing phases developed in this paper provide a promising approach to solve the above mentioned problem although so far only medium size problems are reasonably well-solved. Thus, this work opens interesting possibilities to study-and-develop ad hoc solution procedures that allow us to consider larger size instances of this problem. Moreover, it also points out the possibility of developing heuristic approaches that will give good solutions in competitive running times (Kratka and Stanimirovic, 2006; Puerto et al., 2014). All in all, this paper shows the usefulness of using covering formulations and their corresponding strengthening for solving capacitated versions of ordered hub location problems.

Acknowledgements

This research has been partially supported by Spanish Ministry of Education and Science/FEDER grants numbers MTM2010-19576-C02-(01-02) and MTM2013-46962-C02-(01-02), Junta de Andalucía Grant No. FQM 05849 and Fundación Séneca, Grant No. 08716/PI/08.

References

- Alumur, S., Kara, B.Y., 2008. Network hub location problems: the state of the art. *Euro. J. Oper. Res.* 190 (1), 1–21.
- Aykin, T., 1994. Lagrangian relaxation based approaches to capacitated hub-and-spoke network design problem. *Euro. J. Oper. Res.* 79 (3), 501–523.
- Boland, N., Krishnamoorthy, M., Ernst, A., Ebery, J., 2004. Preprocessing and cutting for multiple allocation hub location problems. *Euro. J. Oper. Res.* 155 (3), 638–653.
- Boland, N., Domínguez-Marín, P., Nickel, S., Puerto, J., 2006. Exact procedures for solving the discrete ordered median problem. *Comput. Oper. Res.* 33, 3270–3300.
- Bollapragada, R., Li, Y., Rao, U.S., 2006. Budget-constrained, capacitated hub location to maximize expected demand coverage in fixed-wireless telecommunication networks. *INFORMS J. Comput.* 18 (4), 422–432.
- Campbell, J.F., 1994. Integer programming formulations of discrete hub location problems. *Euro. J. Oper. Res.* 72, 387–405.
- Campbell, J.F., 1996. Hub location and the p-hub median problem. *Oper. Res.* 44 (6), 923–935.
- Campbell, J.F., O'Kelly, M.E., 2012. Twenty five years of hub location research. *Transport. Sci.* 46 (2), 153–169.
- Campbell, J.F., Ernst, A., Krishnamoorthy, M., 2002. *Hub Location Problems*. Facility Location. Springer, Berlin, pp. 373–407.
- Campbell, A.M., Lowe, T.J., Zhang, L., 2007. The p-hub center allocation problem. *Euro. J. Oper. Res.* 176 (2), 819–835.
- Cánovas, L., García, S., Marín, A., 2007. Solving the uncapacitated multiple allocation hub location problem by means of a dual-ascent technique. *Euro. J. Oper. Res.* 179 (3), 990–1007.
- Contreras, L., Díaz, J.A., Fernández, E., 2009. Lagrangean relaxation for the capacitated hub location problem with single assignment. *OR Spectrum* 31 (3), 483–505.
- Contreras, L., Cordeau, J.F., Laporte, G., 2011. Benders decomposition for large-scale uncapacitated hub location. *Oper. Res.* 59, 1477–1490.
- Correia, I., Nickel, S., Saldanha-da-Gama, F., 2010a. The capacitated single-allocation hub location problem revisited: a note on a classical formulation. *Euro. J. Oper. Res.* 207 (1), 92–96.
- Correia, I., Nickel, S., Saldanha-da-Gama, F., 2010b. Single-assignment hub location problems with multiple capacity levels. *Transport. Res. Part B* 44, 1047–1066.
- De Camargo, R.S., Miranda, G.J., Ferreira, R.P.M., Luna, H.P., 2009. Multiple allocation hub-and-spoke network design under hub congestion. *Comput. Oper. Res.* 36 (12), 3097–3106.
- Ebery, J., Krishnamoorthy, M., Ernst, A., Boland, N., 2000. The capacitated multiple allocation hub location problem: formulations and algorithms. *Euro. J. Oper. Res.* 120 (3), 614–631.
- Elhedhli, S., Hu, F., 2004. Hub-and-spoke network design with congestion. *Comput. Oper. Res.* 32 (6), 1615–1632.
- Ernst, A.T., Krishnamoorthy, M., 1996. Efficient algorithms for the uncapacitated single allocation p-hub median problem. *Location Sci.* 4 (3), 139–154.
- Ernst, A.T., Krishnamoorthy, M., 1999. Solution algorithms for the capacitated single allocation hub location problem. *Ann. Oper. Res.* 86, 141–159.
- Fonseca, M.C., García-Sánchez, A., Ortega-Mier, M., Saldanha-da-Gama, F., 2010. A stochastic bi-objective location model for strategic reverse logistics. *TOP* 18 (1), 158–184.
- García, S., Landete, M., Marín, A., 2012. New formulation and a branch-and-cut algorithm for the multiple allocation p-hub median problem. *Euro. J. Oper. Res.* 220, 48–57.
- Kalcsics, J., Nickel, S., Puerto, J., Rodríguez-Chía, A.M., 2010a. Distribution systems design with role dependent objectives. *Euro. J. Oper. Res.* 202, 491–501.
- Kalcsics, J., Nickel, S., Puerto, J., Rodríguez-Chía, A.M., 2010b. The ordered capacitated facility location problem. *TOP* 18 (1), 203–222.
- Kara, B.Y., Tansel, B.C., 2000. On the single-assignment p-hub center problem. *Euro. J. Oper. Res.* 125 (3), 648–655.
- Kara, B.Y., Tansel, B.C., 2003. The single-assignment hub covering problem: models and linearizations. *J. Oper. Res. Soc.* 54, 59–64.
- Kratika, J., Stanimirovic, Z., 2006. Solving the uncapacitated multiple allocation p-hub center problem by genetic algorithm. *Asia-Pacific J. Oper. Res.* 23 (4), 425–437.
- Labbé, M., Yaman, H., Gourdin, E., 2005. A branch and cut algorithm for hub location problems with single assignment. *Math. Programm.* 102 (2), 371–405.
- 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. Global Optim.* 33 (3), 393–422.
- Marín, A., Cánovas, L., Landete, M., 2006. New formulations for the uncapacitated multiple allocation hub location problem. *Euro. J. Oper. Res.* 172 (1), 274–292.
- Marín, A., Nickel, S., Puerto, J., Velten, S., 2009. A flexible model and efficient solution strategies for discrete location problems. *Discrete Appl. Math.* 157 (5), 1128–1145.
- Marín, A., Nickel, S., Velten, S., 2010. An extended covering model for flexible discrete and equity location problems. *Math. Methods Oper. Res.* 71 (1), 125–163.
- Meyer, T., Ernst, A.T., Krishnamoorthy, M., 2009. A 2-phase algorithm for solving the single allocation p-hub center problem. *Comput. Oper. Res.* 36 (12), 3143–3151.
- Nickel, S., Puerto, J., 2005. *Location Theory – A Unified Approach*. Springer.
- Puerto, J., Ramos, A.B., Rodríguez-Chía, A.M., 2011. Single-allocation ordered median hub location problems. *Comput. Oper. Res.* 38, 559–570.
- Puerto, J., Ramos, A.B., Rodríguez-Chía, A.M., 2013. A specialized Branch & Bound & Cut for single-allocation ordered median hub location problems. *Discrete Appl. Math.* 161, 2624–2646.
- Puerto, J., Pérez-Brito, D., García-González, C.G., 2014. A modified variable neighborhood search for the discrete ordered median problem. *Euro. J. Oper. Res.* 234 (1), 61–76.
- Ramos, A.B., 2012. *Hub Location Problems with Ordered Median (In Spanish)*. Ph.D. Thesis. Facultad de Matemáticas, Universidad de Sevilla.
- Skorin-Kapov, D., Skorin-Kapov, J., O'Kelly, M., 1996. Tight linear programming relaxations of uncapacitated p-hub median problems. *Euro. J. Oper. Res.* 94, 582–593.
- Tan, P.Z., Kara, B.Y., 2007. A hub covering model for cargo delivery systems. *Networks* 49 (1), 28–39.
- Wagner, B., 2008. Model formulations for hub covering problems. *J. Oper. Res. Soc.* 59 (7), 932–938.