Computers & Operations Research 50 (2014) 161–174

Contents lists available at ScienceDirect

Computers & Operations Research

journal homepage: www.elsevier.com/locate/caor

A branch-and-cut algorithm for the hub location and routing problem

Inmaculada Rodríguez-Martín [a,*], Juan-José Salazar-González [a], Hande Yaman [b]

[a] DEIOC, Facultad de Matemáticas, Universidad de La Laguna, Tenerife, Spain
[b] Department of Industrial Engineering, Bilkent University, Ankara, Turkey

ARTICLE INFO

Available online 9 May 2014

Keywords:
Hub location
Routing
Valid inequalities
Branch-and-cut

ABSTRACT

We study the hub location and routing problem where we decide on the location of hubs, the allocation of nodes to hubs, and the routing among the nodes allocated to the same hubs, with the aim of minimizing the total transportation cost. Each hub has one vehicle that visits all the nodes assigned to it on a cycle. We propose a mixed integer programming formulation for this problem and strengthen it with valid inequalities. We devise separation routines for these inequalities and develop a branch-and-cut algorithm which is tested on CAB and AP instances from the literature. The results show that the formulation is strong and the branch-and-cut algorithm is able to solve instances with up to 50 nodes.

1. Introduction

In a network where traffic is collected from many origins to be distributed to many destinations, connecting all origins and destinations with direct links is often not justified in economical terms. Hubbing is used to combine traffic demands from many origins to many destinations and route them together.

The classical *Single Allocation p-Hub Median Problem* (SApHMP) is defined as follows. Let us consider a set of nodes, pairwise traffic demands, routing costs and economies of scale factor. The problem is to select p nodes (called *hubs*) and assign each node to exactly one of these hubs to minimize the total cost of routing the traffic. The traffic from i to j must traverse at least one and at most two hub nodes. (The traffic can go directly from one node to another when one of them is a hub and the other is assigned to it.) If node i is assigned to hub j and node m is assigned to hub l, then the traffic from node i to node m follows the path $i \rightarrow j \rightarrow l \rightarrow m$. Hence the traffic traveling from hub node j to hub node l is the traffic from nodes assigned to hub j to nodes assigned to hub l. This traffic is routed through the hub network at a discounted cost due to economies of scale.

In this study, we consider the *Hub Location and Routing Problem* (HLRP). As in SApHMP, we are given a set of nodes, pairwise traffic demands and routing costs. HLRP consists of selecting p hubs, assigning each node to exactly one of these hubs, and connecting the nodes assigned to each hub with a cycle. Each cycle is limited to at most q nodes. The hub nodes are directly connected by (uncapacitated) links. The aim of the problem is to minimize the total cost of assigning nodes to hubs and the cost of routing the traffic in the network. The traffic between nodes assigned to the same hub is routed on the cycle incident at this hub, whereas the traffic between nodes assigned to different hubs is routed through the hub network and through the cycles. The cost of routing on the cycles is independent of the traffic and is a function of the distance traversed. On the other hand, the routing cost in the hub network is a function of the distance and the traffic. Fig. 1 illustrates a potential HLRP solution for an instance with 4 hubs and 11 non-hub nodes. The solid lines represent the inter-hub complete network. Hubs 1 and 4 have two or more non-hub nodes assigned that are connected to them by a cycle. Hub 2 has only one non-hub node assigned, and hub 3 has none. Then, the traffic going from hub 2 to hub 3 is the sum of the traffic originating at nodes 8 and 2 with destination to node 3. The traffic going from hub 3 to 4 is the sum of the traffic with origin at node 3 and destination to node 4 or to any of the non-hub nodes assigned to it.

HLRP arises in transportation and logistics applications where hubbing is used and nodes do not have sufficient demand to justify direct connections with the hubs (see, e.g. [8,11,18,36,44] for similar situations). In particular, this situation often appears in postal delivery and cargo delivery applications, where many small branch offices are located in population centers and vehicles collect their traffic and carry them to a hub. One of the largest cargo delivery companies in Turkey operates 844 branch offices, most of which are small in traffic volume. Instead of connecting directly each branch office with a hub, which would be very expensive, hubs have vehicles that collect the parcels from the branch offices and bring them to the hub to be sorted and rerouted.

HLRP is a combination of hub location and multi-depot vehicle routing problems and, consequently, it is a difficult problem. The

* Corresponding author.
E-mail addresses: irguez@ull.es (I. Rodríguez-Martín), jjsalaza@ull.es (J.-J. Salazar-González), hyaman@bilkent.edu.tr (H. Yaman).

http://dx.doi.org/10.1016/j.cor.2014.04.014
0305-0548/

literature on HLRP is very limited. Here we first briefly review exact approaches for related problems.

If the costs associated with the cycles are zero, then HLRP reduces to SApHMP. O'Kelly [38] models the SApHMP as a quadratic 0-1 problem. Campbell [12] and Skorin-Kapov et al. [43] propose 4-index linearizations. A 3-index linearization for the special case where the routing costs satisfy the triangle inequality is given by Ernst and Krishnamoorthy [22]. Ebery [21] presents a 2-index linearization and a formulation for two or three hubs. Ernst and Krishnamoorthy [23] propose a branch-and-bound method, where shortest-path problems are solved to compute lower bounds. Labbé and Yaman [28] compare two multicommodity formulations and study their projections. Labbé et al. [30] propose another 2-index formulation, derive valid inequalities and use them in a branch-and-cut algorithm.

For a more extensive review of the studies on SApHMP, we refer the reader to the surveys by Campbell et al. [13], Alumur and Kara [2], and the recent article by Campbell and O'Kelly [16]. Most of these studies are based on many assumptions, such as the hub network is complete, no fixed costs are incurred for routing, each node is connected to a hub with a direct link, and the traffic cost on a hub link is discounted by a factor that does not depend on the amount of flow. In recent years, there have been quite a number of studies trying to relax these assumptions to make the problem more realistic. O'Kelly and Miller [40], Nickel et al. [37], Yoon and Current [50], Calik et al. [9] and Alumur et al. [3] consider hub location problems where the hub network is not necessarily complete. Labbé and Yaman [29], Yaman [46] and Yaman and Elloumi [48] consider star hub networks. Contreras et al. [20] study a tree structure and Yaman [47] and Alumur et al. [4] consider hierarchical hub networks. Campbell et al. [14,15] study the problem of locating a given number of hub arcs with discounted costs rather than locating hubs. Podnar et al. [41] discount the transportation cost on a link if the flow on this link exceeds a threshold. O'Kelly and Brian [39], Horner and O'Kelly [26] and Camargo et al. [10] relax the assumption of a fixed discount factor on hub links and model economies of scale as a function of flow. Yaman et al. [49] study the problem with stopovers with the aim of minimizing the longest travel time and Yaman [45] studies the r-allocation variant where a node can be allocated to up to r hub nodes. Recent studies are mostly focused on relaxing assumptions related to the hub networks. There are few studies on the design of the networks connecting a hub and the nodes assigned to it. We aim to fill this gap in the hub location literature.

If the cost of routing traffic on the hub network is zero, then HLRP reduces to a variant of the plant-cycle location problem for which Labbé et al. [27] propose a branch-and-cut algorithm. Albareda-Sambola et al. [1] propose a compact formulation defined on an auxiliary network and derive lower bounds. The plant-cycle location problem is a special case of location-routing problem where each facility has one vehicle. In the general

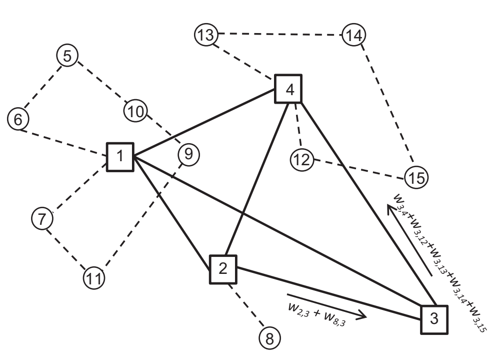

Fig. 1. A HCLP solution example.

location-routing problem, a facility can serve its clients using multiple vehicles. The single facility version of this problem is studied by Laporte and Nobert [31]. Laporte et al. [32,34] propose exact methods to solve the multiple facility problem with capacitated vehicles and maximum route costs, respectively, and Belenguer et al. [7] present a branch-and-cut algorithm.

Another closely related problem is the multi-depot vehicle routing problem (MDVRP). If the hub locations and the number of cycles incident to hubs are fixed and the routing costs on the hub-to-hub links are zero, then HLRP is a MDVRP. There are few studies on exact methods for this problem. Laporte et al. [33,34] propose branch-and-bound algorithms. Baldacci and Mingozzi [6] note that MDVRP is a heterogeneous VRP where the vehicles at each depot are seen as different types of vehicles. They propose an exact algorithm for the heterogeneous VRP and present computational results for MDVRP.

Finally we mention related studies on hub location and routing problems. Nagy and Salhi [36] consider a hub location and routing problem with capacity and distance constraints. The objective function is the sum of the fixed costs of installing hubs and the fixed costs on hub-to-hub links and on routes visiting customers. A customer can be visited by two routes, one for pickup and one for delivery. The authors present a model and propose a nested solution methodology. Çetiner et al. [18] study a multiple allocation hub location and routing problem for the Turkish postal services. They assume that the demand nodes allocated to a hub are served by uncapacitated vehicles that start and end their trips at the hub node. Their problem has two objectives, the minimization of the variable transportation cost and of the number of vehicles needed to achieve a given service level. They minimize the first objective by imposing an upper bound on the number of vehicles. They propose an iterative hubbing and routing heuristic and present computational results using Turkish data where they allow tours of at most 450 km (one day travel time). Camargo et al. [11] study the single allocation version of a similar problem where the lengths of cycles are bounded from above to ensure service quality. They propose a solution approach based on Benders' decomposition. Wasner and Zäpfel [44] study another postal service application where they allow direct connections between non-hub nodes and the routing costs depend on the number of vehicles required. The authors propose a heuristic method and present a case study using data from Austria. Different from the studies above, in HLRP, each hub has a single vehicle and each vehicle can service at most q nodes. We summarize the different features of the related studies in Table 1. Some other variants of hub location and routing problems have been addressed by Aykin [5], Catanzaro et al. [17], and Rieck et al. [42].

This paper proposes strong formulations for HLRP and describes an exact solution method. Our study contributes to the literature by proposing a solution methodology that handles decisions on different levels of the network simultaneously to find an optimal solution. The rest of the paper is organized as follows. Section 2 presents the notation, a mixed integer programming formulation and valid inequalities. Section 3 describes a branch-and-cut algorithm. We present the results of our computational experiments in Section 4 and write conclusions in Section 5.

2. MIP formulation and valid inequalities

We first introduce the notation. Let V be the set of nodes and p be the number of hubs to open. We denote the traffic demand from node $i \in V$ to node $m \in V$ by w_{im} and the cost of routing a unit of traffic from node $j \in V$ to node $l \in V$ by c_{jl}. We assume that the routing costs satisfy the triangle inequality. Let $o_i = \sum_{m \in V} w_{im}$ be the total amount of demand originating at node i and

Table 1
Studies on hub location and routing.

Study	Allocation	Num. of hubs	Objectives	Capacity	Cycle length bound	Num. of vehicles	Solution approach
Nagy and Salhi [36]	One for pickup One for delivery	Not fixed	Cost	Yes	Yes	Not fixed	MIP formulation, heuristic
Çetiner et al. [18]	Multiple	Not fixed	Cost and num. of vehicles	No	Yes	Not fixed	Heuristic
Camargo et al. [11]	Single	Not fixed	Cost	No	Yes	Not fixed	MIP formulation, Benders decomposition
Wasner and Zäpfel [44]	Multiple direct shipment	Not fixed	Cost	Yes	Yes	Not fixed	MIP formulation, heuristic
Present study	Single	Fixed to p	Cost	At most q nodes per cycle	No	One per hub	MIP formulation, branch-and-cut

$d_i = \sum_{m \in V} w_{mi}$ be the total amount of demand with destination at node i. We compute the cost of assigning node i to hub j as $c_{ij}o_i + c_{ji}d_i$. When i and j are hubs then the routing cost is reduced by a factor of α. We can assign at most $q \geq 2$ nodes (including the hub itself) to a hub. Clearly, we need $pq \geq |V|$ for feasibility.

Let $G=(V,E)$ be an undirected graph, with $E=\{[i,j] : i,j \in V, i \neq j\}$, representing the links that can be in a cycle. For $S \subseteq V$, let $\delta(S)$ be the set of edges with exactly one endpoint in S and $E(S)$ be the set of edges with both endpoints in set S. If $S=\{i\}$, we simply write $\delta(i)$ instead of $\delta(S)$. We denote by f_e the cost of using the edge $e \in E$ in a cycle. A cycle can be defined by two nodes i and j, thus an edge $[i,j]$ can be used twice, and then the cost of the cycle is $2f_{ij}$. We use a factor β to change the relative weight of the cycle edge costs in the objective function. Note that β is not related to the discount factor for collection used in the classical hub location models.

We define z^1_{jj} to be 1 if node $j \in V$ is a hub and no other node is assigned to j, z^1_{ij} to be 1 if node $j \in V$ is a hub and node $i \in V \backslash \{j\}$ is the only other node assigned to it, and z^2_{ij} to be 1 if node $j \in V$ is a hub with at least two other nodes assigned to it and if node $i \in V$ is assigned to j. The variables take value 0 otherwise. With these definitions, node j is a hub if $\sum_{i \in V} z^1_{ij} + z^2_{jj}$ is 1. Node i is assigned to node $j \neq i$ if $z^1_{ij} + z^2_{ij} = 1$. The flow variable g^i_{jl} represents the amount of traffic that originates at node $i \in V$ and travels from hub $j \in V$ to hub $l \in V \backslash \{j\}$. We also use the edge variable x_e for each $e \in E$ to represent the cycles with at least three edges. For $E' \subseteq E$, we define $x(E') := \sum_{e \in E'} x_e$.

The HLRP can be modeled as follows:

$$\min \sum_{i \in V} \sum_{j \in V \backslash \{i\}} (c_{ij}o_i + c_{ji}d_i)(z^1_{ij} + z^2_{ij}) + \alpha \sum_{j \in V} \sum_{l \in V \backslash \{j\}} c_{jl} \sum_{i \in V} g^i_{jl} + \beta \left(\sum_{i \in V} \sum_{j \in V \backslash \{i\}} 2 f_{ij} z^1_{ij} + \sum_{e \in E} f_e x_e \right) \quad (1)$$

$$\text{s.t.} \quad \sum_{j \in V \backslash \{i\}} z^1_{ij} + z^1_{ii} + \sum_{j \in V \backslash \{i\}} z^1_{ji} + \sum_{j \in V} z^2_{ij} = 1 \quad \forall i \in V, \quad (2)$$

$$\sum_{i \in V} z^2_{ij} \leq q z^2_{jj} \quad \forall j \in V, \quad (3)$$

$$\sum_{j \in V} \left(\sum_{i \in V} z^1_{ij} + z^2_{jj} \right) = p, \quad (4)$$

$$\sum_{l \in V \backslash \{j\}} g^i_{jl} - \sum_{l \in V \backslash \{j\}} g^i_{lj} = \sum_{m \in V \backslash \{i,j\}} w_{im}(z^1_{ij} + z^2_{ij} - z^1_{mj} - z^2_{mj}) + w_{ij} \left(z^1_{ij} + z^2_{ij} - \sum_{k \in V} z^1_{kj} - z^2_{jj} \right) \quad \forall i,j \in V, \ i \neq j, \quad (5)$$

$$\sum_{l \in V \backslash \{j\}} g^j_{jl} - \sum_{l \in V \backslash \{j\}} g^j_{lj} = \sum_{m \in V \backslash \{j\}} w_{jm} \left(\sum_{k \in V} z^1_{kj} + z^2_{jj} - z^1_{mj} - z^2_{mj} \right) \quad \forall j \in V, \quad (6)$$

$$x(\delta(i)) = 2 \sum_{j \in V} z^2_{ij} \quad \forall i \in V, \quad (7)$$

$$x(\delta(S)) \geq 2 \sum_{j \in V \backslash S} z^2_{ij} \quad \forall S \subset V, \ i \in S, \quad (8)$$

$$x_{ii'} + z^2_{ij} + z^2_{i'j'} \leq 2 \quad \forall [i,i'] \in E, \ j,j' \in V, \ j \neq j', \quad (9)$$

$$x_e \in \{0,1\} \quad \forall e \in E, \quad (10)$$

$$z^1_{ij} \in \{0,1\} \quad \forall i,j \in V, \quad (11)$$

$$z^2_{ij} \in \{0,1\} \quad \forall i,j \in V, \quad (12)$$

$$g^i_{jl} \geq 0 \quad \forall i,j \in V, \ l \in V \backslash \{j\}. \quad (13)$$

The objective function (1) is the sum of the classical objective function in hub location problems (i.e., the cost of assigning nodes to hubs and the cost of routing in the hub network) plus new terms to consider the routing part within a cycle (i.e., the cost of cycles of two edges and the cost of cycles of at least three edges). Constraints (2) impose that a node i is either the only node assigned to another hub (case $\sum_{j \in V \backslash \{i\}} z^1_{ij} = 1$), or it is a hub with no other node assigned to it (case $z^1_{ii} = 1$), or it is a hub with one other node assigned to it (case $\sum_{j \in V \backslash \{i\}} z^1_{ji} = 1$), or it is a hub or is assigned to another hub with at least two other nodes (case $\sum_{j \in V} z^2_{ij} = 1$). Constraints (3) are capacity constraints to guarantee that a cycle does not contain more than q nodes. The number of hubs to open is p due to constraint (4). Constraints (5) and (6) are the flow balance constraints for the traffic on the hub network. Constraints (7) state that two edges should be adjacent to a node that is assigned to a hub with at least two other nodes. Constraints (8) ensure the connectivity of the cycles. If a node $i \in S$ is assigned to a hub in set $V \backslash S$, then the cycle that contains node i has to cross the cut defined by subset S and $x(\delta(S)) \geq 2$. Constraints (9) forbid nodes assigned to different hubs to be on the same cycle. Constraints (10)–(13) are variable restrictions. Note that (1)–(13) is a model for the SApHMP when $\beta = 0$, and for the plant cycle location problem when $\alpha = 0$ and the constraints involving flow variables g^i_{jl} are excluded.

In the remaining part of this section, we provide several families of valid inequalities. The first family is

$$z^2_{ij} \leq z^2_{jj} \quad \forall i,j \in V, \ i \neq j. \quad (14)$$

Similar inequalities are

$$x_{ii'} + z^2_{i'i'} \leq 1 + z^2_{ii'} \quad \forall [i,i'] \in E, \quad (15)$$

stating that if edge $[i,i']$ is part of a cycle and node i' is a hub, then node i is assigned to hub i'.

Next, we propose a family of valid inequalities that dominate both constraints (9) and the valid inequalities (15).

Table 2
Results for CAB25 with $q=|V|$.

p	α	β	r-gap	r-time	nodes	time	nCuts	totalCost	%-access	%-interHub	%-cycle	hubs
3	0.2	0.01	98.76	2.85	90	4.40	424	858.76	74.04	16.22	9.74	5; 12; 17
		0.05	100.00	2.17	0	2.17	395	1193.41	53.28	11.67	35.05	5; 12; 17
		0.2	100.00	2.28	0	2.29	440	2448.35	25.97	5.69	68.34	5; 12; 17
	0.4	0.01	97.40	3.03	73	4.77	402	998.04	63.71	27.91	8.38	5; 12; 17
		0.05	99.48	3.78	6	4.51	433	1332.69	47.71	20.90	31.39	5; 12; 17
		0.2	100.00	4.59	0	4.60	533	2587.63	24.57	10.77	64.66	5; 12; 17
	0.8	0.01	96.10	4.73	292	10.97	461	1254.02	52.45	39.96	7.59	2; 4; 12
		0.05	97.00	6.72	492	27.30	588	1605.91	50.25	23.70	26.05	5; 8; 18
		0.2	98.89	9.13	15	11.28	593	2827.03	31.29	13.61	55.10	12; 20; 23
4	0.2	0.01	99.02	1.61	29	3.40	288	720.84	66.08	22.74	11.17	4; 12; 14; 17
		0.05	100.00	1.72	0	1.72	288	1041.09	46.17	15.86	37.97	4; 12; 14; 17
		0.2	100.00	2.17	0	2.18	399	2227.04	21.58	7.42	71.00	4; 12; 14; 17
	0.4	0.01	98.34	2.25	30	3.93	304	876.30	55.25	34.62	10.13	1; 4; 12; 17
		0.05	99.30	2.25	13	3.78	327	1206.25	39.84	27.38	32.77	4; 12; 14; 17
		0.2	99.36	5.01	14	5.88	506	2392.19	20.09	13.81	66.10	4; 12; 14; 17
	0.8	0.01	96.06	3.43	721	18.19	476	1176.44	42.63	49.83	7.55	1; 4; 12; 18
		0.05	95.87	5.52	1190	30.17	639	1528.42	43.11	30.54	26.35	4; 8; 18; 24
		0.2	99.44	9.14	15	10.87	618	2615.26	31.40	15.47	53.13	8; 12; 20; 23
5	0.2	0.01	99.62	2.23	12	2.73	235	626.71	59.02	26.95	14.03	4; 7; 12; 14; 17
		0.05	100.00	2.07	0	3.53	279	947.54	40.40	18.31	41.29	4; 7; 12; 14; 17
		0.2	100.00	2.70	0	2.95	442	2027.18	22.00	8.00	70.00	4; 12; 14; 17; 23
	0.4	0.01	98.66	1.76	61	4.15	223	795.61	46.49	42.46	11.05	4; 7; 12; 14; 17
		0.05	98.52	2.34	64	4.99	314	1120.99	34.15	30.95	34.90	4; 7; 12; 14; 17
		0.2	99.67	3.79	4	4.59	455	2179.65	25.55	12.83	61.62	5; 8; 12; 17; 23
	0.8	0.01	95.60	6.40	2913	47.11	741	1126.18	37.58	54.24	8.18	1; 4; 7; 12; 18
		0.05	96.24	8.36	827	25.80	481	1446.56	32.73	42.23	25.04	4; 12; 18; 23; 24
		0.2	98.76	6.19	74	9.31	560	2457.77	31.93	16.95	51.12	8; 12; 20; 22; 23

Table 3
Results for CAB25 with $q=\lceil |V|/2 \rceil$.

p	α	β	r-gap	r-time	nodes	time	nCuts	totalCost	%-access	%-interHub	%-cycle	hubs
3	0.2	0.01	98.35	4.84	128	6.60	407	865.42	72.58	16.49	10.93	4; 12; 17
		0.05	99.03	3.31	24	4.01	436	1213.10	52.08	12.47	35.45	12; 17; 21
		0.2	99.39	5.68	15	7.13	594	2495.76	25.45	6.12	68.43	12; 17; 21
	0.4	0.01	97.52	3.35	67	4.68	388	999.62	64.16	26.37	9.47	4; 12; 18
		0.05	97.77	4.99	398	15.01	588	1359.94	47.87	20.51	31.62	5; 12; 17
		0.2	98.30	7.33	322	16.65	585	2648.59	23.98	11.54	64.48	12; 17; 21
	0.8	0.01	96.05	5.40	601	20.87	683	1254.02	52.45	39.96	7.59	2; 4; 12
		0.05	96.06	5.07	558	21.64	641	1623.26	41.34	30.51	28.15	12; 21; 25
		0.2	97.05	14.43	1103	56.69	983	2917.72	27.98	13.07	58.95	5; 8; 18
4	0.2	0.01	99.06	1.64	26	2.14	255	720.84	66.08	22.74	11.17	4; 12; 14; 17
		0.05	100.00	1.54	0	1.54	268	1041.09	46.17	15.86	37.97	4; 12; 14; 17
		0.2	100.00	1.36	0	1.42	339	2227.04	21.58	7.42	71.00	4; 12; 14; 17
	0.4	0.01	98.26	2.00	39	4.17	283	876.30	55.25	34.62	10.13	1; 4; 12; 17
		0.05	99.28	2.56	17	3.20	336	1206.25	39.84	27.38	32.77	4; 12; 14; 17
		0.2	99.59	3.71	24	4.48	479	2392.19	20.09	13.81	66.10	4; 12; 14; 17
	0.8	0.01	96.08	4.96	889	23.07	512	1176.44	42.63	49.83	7.55	1; 4; 12; 18
		0.05	95.97	5.85	1736	53.17	667	1528.42	43.11	30.54	26.35	4; 8; 18; 24
		0.2	96.72	11.22	535	29.14	797	2716.52	23.77	18.02	58.21	4; 8; 17; 24
5	0.2	0.01	99.62	2.03	13	2.43	225	626.71	59.02	26.95	14.03	4; 7; 12; 14; 17
		0.05	99.95	1.51	2	1.67	271	947.54	40.40	18.31	41.29	4; 7; 12; 14; 17
		0.2	99.47	5.18	14	6.29	543	2040.02	21.70	8.27	70.02	4; 12; 14; 17; 23
	0.4	0.01	98.69	2.12	42	3.17	231	795.61	46.49	42.46	11.05	4; 7; 12; 14; 17
		0.05	98.43	2.48	26	3.37	327	1120.99	34.15	30.95	34.90	4; 7; 12; 14; 17
		0.2	98.69	5.82	39	7.25	503	2208.83	20.04	15.28	64.67	4; 12; 14; 17; 23
	0.8	0.01	95.54	4.23	3788	62.45	923	1126.18	37.58	54.24	8.18	1; 4; 7; 12; 18
		0.05	96.28	6.08	881	23.28	568	1446.56	32.73	42.23	25.04	4; 12; 18; 23; 24
		0.2	97.17	6.16	1106	43.21	925	2514.93	23.33	21.41	55.26	5; 8; 12; 18; 23

Proposition 1.

$$x_{ii'} \le \sum_{j \in V \setminus S} z_{ij}^2 + \sum_{j' \in S} z_{i'j'}^2 \tag{16}$$

is valid for all $[i,i'] \in E$ *and* $S \subset V$ *such that* $i \in S$ *and* $i' \in V \setminus S$.

Proof. If $\sum_{j \in V \setminus S} z_{ij}^2 + \sum_{j' \in S} z_{i'j'}^2 = 0$, then $\sum_{j \in V \setminus \{i\}} z_{ij}^1 + z_{ii}^1 + \sum_{j \in V \setminus \{i\}} z_{ji}^1 + \sum_{j \in S} z_{ij}^2 = 1$ and $\sum_{j' \in V \setminus \{i'\}} z_{i'j'}^1 + z_{i'i'}^1 + \sum_{j' \in V \setminus \{i'\}} z_{j'i'}^1 + \sum_{j' \in V \setminus S} z_{i'j'}^2 = 1$. The first part implies that node i is assigned to a hub in set S, or it is in a cycle of length two, or it is a hub node with no other nodes assigned to it. Similarly, the second part implies that node i' is assigned to a hub in set $V \setminus S$, or it is in a cycle of length two, or it is an isolated hub. In all cases, i and i' cannot be in the same cycle and, as a result, edge $[i,i']$ cannot be in a cycle. □

Note that inequalities (16) dominate constraints (9) since

$$x_{ii'} + \sum_{j \in S} z_{ij}^2 + \sum_{j' \in V \setminus S} z_{i'j'}^2 \le 2 - \sum_{k \in V \setminus \{i,i'\}} (z_{ik}^1 + z_{ki}^1 + z_{i'k}^1 + z_{ki'}^1) - z_{ii}^1 - z_{i'i'}^1 - 2z_{ii'}^1 - 2z_{i'i}^1.$$

The particular case of inequalities (16) with $S = V \setminus \{i'\}$ is

$$x_{ii'} \le z_{ii'}^2 + \sum_{j' \in V \setminus \{i'\}} z_{i'j'}^2$$

which is the same as

$$x_{ii'} + z_{i'i'}^2 \le 1 + z_{ii'}^2 - \left(1 - \sum_{j' \in V} z_{i'j'}^2\right).$$

Observe that the above inequality dominates inequality (15) since $1 - \sum_{j' \in V} z_{i'j'}^2 \ge 0$.

The following family of inequalities is used by Labbé et al. [27] to solve the plant cycle location problem. Let $S \subset V$, $i \in S$ and $i' \in V \setminus S$. The generalized subtour elimination constraint is

$$x(\delta(S)) \ge 2\left(\sum_{j \in V \setminus S} z_{ij}^2 + \sum_{j \in S} z_{i'j}^2\right). \tag{17}$$

Inequalities (17) are stronger than constraints (8). If $\sum_{j \in V \setminus S} z_{ij}^2 + \sum_{j \in S} z_{i'j}^2 = 2$, then node $i \in S$ is assigned to a hub in $V \setminus S$ and node $i' \in V \setminus S$ is assigned to a hub in S. Hence, at least two cycles cross the cut and as a result $x(\delta(S)) \ge 4$.

Next, we consider inequalities that take into account the capacity constraints.

Proposition 2.

$$x(E(S)) - \sum_{i \in S} z_{ii}^2 \le |S| - \left\lceil \frac{|S|}{q} \right\rceil \tag{18}$$

is valid for all $S \subseteq V$.

Proof. We first prove that all feasible solutions satisfy

$$x(E(S)) - \sum_{i \in S} z_{ii}^2 \le \sum_{i \in S} \sum_{j \in V} z_{ij}^2 - \left\lceil \frac{\sum_{i \in S} \sum_{j \in V} z_{ij}^2}{q} \right\rceil. \tag{19}$$

When $\sum_{i \in S} z_{ii}^2 = 0$ set S contains no hub node that is on a cycle of at least three nodes. In this case it is easy to see that the number of edges of a solution (that is, edges in cycles) inside the set S cannot be more than

$$\sum_{i \in S} \sum_{j \in V} z_{ij}^2 - \left\lceil \frac{\sum_{i \in S} \sum_{j \in V} z_{ij}^2}{q} \right\rceil.$$

For each hub node $i \in S$ with $z_{ii}^2 = 1$ we may have at most one additional edge of a cycle inside S. Then, the inequality is satisfied by all feasible solutions.

Inequality (18) is valid since

$$|S| - \left\lceil \frac{|S|}{q} \right\rceil \ge \sum_{i \in S} \sum_{j \in V} z_{ij}^2 - \left\lceil \frac{\sum_{i \in S} \sum_{j \in V} z_{ij}^2}{q} \right\rceil. \qquad \square$$

Table 4
Results for CAB25 with $q = \lceil |V|/p \rceil$.

p	α	β	r-gap	r-time	nodes	time	nCuts	totalCost	%-access	%-interHub	%-cycle	hubs
3	0.2	0.01	98.04	3.98	114	5.57	391	943.25	72.81	16.31	10.88	4; 12; 18
		0.05	96.23	6.66	971	26.99	988	1348.93	50.99	11.42	37.59	4; 12; 18
		0.2	95.05	12.01	3611	151.59	2148	2789.59	26.22	6.15	67.63	12; 13; 17
	0.4	0.01	95.76	7.02	391	16.24	865	1089.05	65.49	25.09	9.42	4; 18; 19
		0.05	94.40	8.24	1381	49.23	1529	1494.95	47.78	18.30	33.92	4; 18; 19
		0.2	94.03	12.82	4828	348.46	2650	2926.27	27.87	7.82	64.31	5; 8; 17
	0.8	0.01	95.32	5.49	951	26.35	866	1302.98	60.98	31.24	7.78	2; 4; 8
		0.05	93.98	8.44	5755	230.48	2035	1708.64	46.50	23.82	29.68	2; 4; 8
		0.2	95.17	8.28	3197	208.74	2318	3099.42	32.94	8.45	58.61	1; 2; 4
4	0.2	0.01	99.82	0.62	3	0.80	141	721.98	64.54	23.65	11.81	4; 12; 16; 17
		0.05	99.88	1.20	3	1.25	208	1063.03	43.83	16.06	40.10	4; 12; 16; 17
		0.2	99.85	5.23	3	5.73	573	2341.94	19.90	7.29	72.81	4; 12; 16; 17
	0.4	0.01	99.11	2.31	20	2.89	242	881.26	54.57	35.75	9.67	1; 4; 12; 17
		0.05	99.71	3.99	13	4.84	373	1222.30	39.35	25.78	34.88	1; 4; 12; 17
		0.2	99.46	5.94	8	8.25	509	2501.22	19.23	12.60	68.18	1; 4; 12; 17
	0.8	0.01	96.61	6.22	400	18.42	542	1178.69	42.82	49.58	7.60	1; 4; 12; 18
		0.05	97.32	8.36	648	20.78	525	1531.41	42.37	29.79	27.84	1; 4; 8; 18
		0.2	97.69	12.37	306	25.30	628	2810.33	23.09	16.23	60.68	1; 4; 8; 18
5	0.2	0.01	97.05	4.04	242	7.50	433	686.85	61.14	24.89	13.97	4; 6; 12; 17; 24
		0.05	95.82	4.56	1165	22.53	1016	1050.38	40.80	16.57	42.64	9; 11; 12; 17; 24
		0.2	94.72	4.46	3901	126.97	2202	2393.88	17.90	7.27	74.83	9; 11; 12; 17; 24
	0.4	0.01	95.99	2.79	1134	14.79	466	857.24	52.29	36.52	11.20	1; 4; 6; 12; 17
		0.05	95.31	3.68	1000	16.26	650	1221.48	37.60	25.73	36.66	1; 9; 12; 17; 21
		0.2	94.64	5.44	11106	1011.78	5598	2564.98	17.91	12.25	69.84	1; 9; 12; 17; 21
	0.8	0.01	95.39	4.35	4103	107.34	1786	1165.79	51.37	40.40	8.23	1; 4; 6; 8; 17
		0.05	95.29	5.97	2176	63.84	2128	1532.11	39.81	30.96	29.23	1; 8; 9; 17; 21
		0.2	94.72	7.22	4094	165.03	2588	2875.61	21.21	16.49	62.29	1; 8; 9; 17; 21

Using Eqs. (7), inequalities (19) can be written as

$$x(\delta(S)) \geq 2\left(\left\lceil \frac{\sum_{i\in S}\sum_{j\in V} z_{ij}^2}{q}\right\rceil - \sum_{i\in S} z_{ii}^2\right).$$

These inequalities are stronger than (18) but non-linear. Another alternative to obtain linear inequalities is motivated by the Multi-Star inequalities for the Capacitated Vehicle Routing Problem (see Letchford et al. [35]). Indeed, by simply removing the rounding operator we get the *fractional capacity cuts*:

$$x(\delta(S)) \geq 2\left(\frac{\sum_{i\in S}\sum_{j\in V} z_{ij}^2}{q} - \sum_{i\in S} z_{ii}^2\right),$$

which can be strengthened as follows.

Proposition 3.

$$x(\delta(S)) \geq 2\left(\frac{\sum_{i\in S}\sum_{j\in V} z_{ij}^2 + x(\delta(S))/2}{q} - \sum_{i\in S} z_{ii}^2\right) \tag{20}$$

is valid for all $S \subseteq V$.

Proof. A vehicle serving nodes in S must also have capacity to serve the nodes not in S visited immediately after serving a node in S. The number of such nodes is $x(\delta(S))$ in general but, since there may be only one node in the vehicles's route outside S, that figure must be divided by two. □

3. Branch-and-cut algorithm

In this section we describe a branch-and-cut algorithm to solve the HLRP. The branch-and-cut scheme for integer programming problems combines a branch-and-bound method for exploring a decision tree and a cutting plane method for computing bounds. At each node of the search tree, the cutting plane method improves a linear relaxation of the problem. When this is not further possible, the branch-and-bound algorithm proceeds. A key point is to have a mathematical model whose linear relaxation is close to the integer problem, and efficient procedures to solve the separation problems and identify violated inequalities within the cutting plane phase.

We now outline the main features of our branch-and-cut algorithm.

3.1. Initial relaxation

At the root node of the branch-and-cut tree we initialize the linear program (LP) model by relaxing constraints (8) and (9) as well as the integrality constraints on the variables of the original

Table 5
Results for AP25.

q	p	β	r-gap	r-time	nodes	time	nCuts	totalCost	%-access	%-interHub	%-cycle	hubs
$\|V\|$	3	1	99.42	3.34	3	3.74	291	155,482.14	85.53	14.32	0.15	7; 14; 18
		100	99.13	4.17	39	6.30	315	177,838.26	74.78	12.52	12.70	7; 14; 18
		500	99.54	7.82	18	10.84	349	262,544.57	50.08	10.27	39.65	2; 9; 18
		1000	99.75	11.14	22	15.91	439	366,638.05	35.86	7.36	56.78	2; 9; 18
	4	1	98.98	3.15	57	5.74	270	139,430.10	79.91	19.93	0.17	2; 7; 14; 18
		100	98.73	4.65	129	11.76	269	161,485.26	69.08	18.29	12.62	2; 9; 17; 18
		500	99.85	6.49	3	8.03	329	243,004.56	45.91	12.16	41.93	2; 9; 17; 18
		1000	100.00	9.94	0	10.72	376	344,903.68	32.35	8.57	59.09	2; 9; 17; 18
	5	1	99.52	2.54	12	3.35	215	123,802.90	76.37	23.45	0.18	2; 7; 14; 17; 18
		100	99.86	4.99	10	6.18	263	145,099.06	64.51	20.76	14.73	2; 8; 17; 18; 20
		500	99.90	4.68	11	6.13	275	227,204.68	43.74	12.32	43.94	2; 7; 14; 17; 18
		1000	100.00	3.42	0	3.60	259	327,043.26	30.39	8.56	61.06	2; 7; 14; 17; 18
$\lceil \|V\|/2 \rceil$	3	1	99.44	3.09	3	3.68	292	155,482.14	85.53	14.32	0.15	7; 14; 18
		100	99.26	5.76	7	6.52	306	177,838.26	74.78	12.52	12.70	7; 14; 18
		500	99.31	5.96	13	8.13	310	262,544.57	50.08	10.27	39.65	2; 9; 18
		1000	99.55	8.02	36	11.11	390	366,638.05	35.86	7.36	56.78	2; 9; 18
	4	1	98.94	2.89	35	5.24	266	139,430.10	79.91	19.93	0.17	2; 7; 14; 18
		100	98.65	3.74	87	7.96	250	161,485.26	69.08	18.29	12.62	2; 9; 17; 18
		500	99.92	5.41	6	6.61	316	243,004.56	45.91	12.16	41.93	2; 9; 17; 18
		1000	100.00	8.13	0	8.28	352	344,903.68	32.35	8.57	59.09	2; 9; 17; 18
	5	1	99.57	2.53	13	3.48	209	123,802.90	76.37	23.45	0.18	2; 7; 14; 17; 18
		100	99.93	3.53	5	4.68	243	145,099.06	64.51	20.76	14.73	2; 8; 17; 18; 20
		500	99.91	4.77	3	5.77	256	227,204.68	43.74	12.32	43.94	2; 7; 14; 17; 18
		1000	100.00	4.12	0	4.26	257	327,043.26	30.39	8.56	61.06	2; 7; 14; 17; 18
$\lceil \|V\|/p \rceil$	3	1	100.00	1.84	0	1.86	230	156,287.34	84.78	15.08	0.14	7; 14; 18
		100	100.00	2.82	0	2.95	241	178,328.06	74.30	13.21	12.48	7; 14; 18
		500	99.14	11.20	96	18.27	384	267,381.48	49.56	8.81	41.63	7; 14; 18
		1000	98.54	12.29	295	32.76	431	376,932.18	37.75	4.98	57.27	8; 17; 18
	4	1	99.47	3.14	5	4.12	208	139,876.23	83.02	16.82	0.16	7; 14; 17; 18
		100	99.50	4.15	5	5.49	219	161,720.99	71.80	14.55	13.64	7; 14; 17; 18
		500	98.26	14.15	193	34.71	417	249,982.62	46.45	9.41	44.13	7; 14; 17; 18
		1000	97.21	12.07	1335	106.14	1726	359,669.90	33.42	6.45	60.13	7; 14; 17; 18
	5	1	99.68	4.82	12	5.63	243	130,727.14	74.88	24.94	0.18	2; 7; 14; 17; 18
		100	98.64	3.92	91	7.05	251	154,151.28	63.50	21.15	15.35	2; 7; 14; 17; 18
		500	97.40	6.38	757	29.69	728	245,105.99	40.65	13.40	45.95	2; 7; 14; 17; 18
		1000	97.52	16.30	2851	166.38	2004	357,731.82	27.85	9.18	62.97	2; 7; 14; 17; 18

formulation. Hence, the initial LP model is (1)–(7) and the continuous relaxation of (10)–(13).

3.2. Cutting plane phase

Given a fractional solution $(x^*, z^{*1}, z^{*2}, g^*)$, the separation routines for constraints (14)–(16), (20), (8), and (18) are applied, in this sequence. The violation of constraints (17) is checked only if no other violated cuts have been found. The number of cuts added to the model in each cut generation step is limited to 100.

3.2.1. Separation of inequalities (8)

The subtour elimination constraints for the TSP can be separated in polynomial time by solving a max-flow/min-cut problem on an appropriately defined support graph. We follow the same idea to devise a separation procedure for inequalities (8), similar to the one used in Labbé et al. [27]. First note that inequalities (8) can be written as

$$x(\delta(S)) + 2\sum_{j \in S} z_{ij}^2 \geq 2\sum_{j \in V} z_{ij}^2 \quad \forall S \subset V, \ i \in S.$$

For each given node $i \in V$ such that $\sum_{j \in V} z_{ij}^{*2} > 0$, let us consider a graph $G' = (V', E')$ with $V' = V \cup \{s\}$, where s is a dummy node. The edge set E' contains all edges $e \in E$ such that $x_e^* > 0$, plus all edges connecting s with nodes $j \in V$ such that $z_{ij}^{*2} > 0$. The capacity of an edge $e \in E$ is x_e^*, and the capacity of an edge $[s,j]$ is $2z_{ij}^{*2}$. Then, a set $S \subset V'$ with $s \notin S$ and $i \in S$ defines a violated inequality (8) if the capacity of the cut $\delta(S)$ is smaller than $2\sum_{j \in V} z_{ij}^{*2}$. Hence, finding the most violated inequality (8) involving i, if any, is equivalent to solving a min-cut problem for i and s on G'.

We solve each min-cut problem using the path-relabel flow algorithm proposed by Goldberg and Tarjan [25], which has complexity $O(mn^2)$ on a graph with n vertices and m edges. So, the overall complexity of our separation procedure is $O(|E||V|^3)$.

3.2.2. Separation of inequalities (14)–(16)

We separate constraints (15), despite being dominated by constrains (16), because they have proven to be useful in our computational experiments. Constraints (14) and (15) can be separated in $O(|V|^2)$ by complete enumeration.

Inequalities (16) can also be separated in polynomial time. For a given edge $[i,i'] \in E$, we define $S = \{i\} \cup \{j \in V \backslash \{i'\} : z_{ij}^{*2} \geq z_{i'j}^{*2}\}$. If the inequality for this choice of S is not violated, then there exists no violated inequality (16) for edge $[i,i']$. The overall complexity of the separation algorithm is $O(|V|^3)$.

3.2.3. Separation of inequalities (17)

The separation procedure for inequalities (17) is an adaptation of the one used in Labbé et al. [27]. We can rewrite constraints (17) as

$$x(\delta(S)) + 2\sum_{j \in S} z_{ij}^2 + 2\sum_{j \in V \backslash S} z_{i'j}^2 \geq 2\left(\sum_{j \in V} z_{ij}^2 + \sum_{j \in V} z_{i'j}^2\right) \quad \forall S \subset V, \ i \in S, \ i' \in V \backslash S.$$

Table 6
Results for AP40.

q	p	β	r-gap	r-time	nodes	time	nCuts	totalCost	%-access	%-interHub	%-cycle	hubs
$\lvert V\rvert$	3	1	99.56	35.94	7	44.94	614	159,131.34	85.63	14.19	0.19	12; 22; 28
		100	99.11	59.09	538	181.37	838	188,910.27	72.13	11.95	15.92	12; 22; 28
		500	98.55	100.01	1128	688.68	1564	306,243.01	44.46	7.98	47.56	12; 23; 28
		1000	99.26	153.57	676	646.09	1583	445,218.18	33.55	5.57	60.88	4; 13; 28
	4	1	99.41	30.03	13	39.73	532	144,269.55	82.46	17.33	0.21	12; 22; 26; 28
		100	99.06	46.11	245	108.81	925	174,036.22	68.36	14.37	17.28	12; 22; 26; 28
		500	98.42	129.76	2910	1322.81	1711	291,653.08	42.96	9.03	48.01	7; 12; 23; 28
		1000	98.64	229.32	960	910.67	1254	430,540.90	29.19	6.65	64.16	4; 12; 15; 28
	5	1	98.85	50.25	413	105.38	913	134,569.34	78.77	21.00	0.23	3; 12; 22; 26; 28
		100	98.54	48.44	1092	411.62	1961	164,038.24	65.62	16.74	17.64	7; 12; 22; 26; 28
		500	98.86	99.12	811	518.69	1422	277,247.49	40.49	11.01	48.50	4; 7; 12; 23; 28
		1000	99.54	126.06	128	186.53	892	411,710.46	27.27	7.41	65.32	4; 7; 12; 23; 28
$\lceil \lvert V\rvert/2 \rceil$	3	1	99.61	34.46	9	45.86	572	159,131.34	85.63	14.19	0.19	12; 22; 28
		100	99.19	61.67	454	233.88	1194	188,910.27	72.13	11.95	15.92	12; 22; 28
		500	98.48	84.79	1121	712.00	1718	306,243.01	44.46	7.98	47.56	12; 23; 28
		1000	99.24	146.75	647	600.46	1852	445,218.18	33.55	5.57	60.88	4; 13; 28
	4	1	99.34	31.48	42	46.22	519	144,269.55	82.46	17.33	0.21	12; 22; 26; 28
		100	99.10	61.28	241	130.92	644	174,036.22	68.36	14.37	17.28	12; 22; 26; 28
		500	98.38	110.42	1178	773.80	1342	291,653.08	42.96	9.03	48.01	7; 12; 23; 28
		1000	98.63	160.09	2386	1900.23	3146	430,540.90	29.19	6.65	64.16	4; 12; 15; 28
	5	1	98.88	42.92	398	138.65	792	134,569.34	78.77	21.00	0.23	3; 12; 22; 26; 28
		100	98.47	47.99	459	133.82	980	164,038.24	65.62	16.74	17.64	7; 12; 22; 26; 28
		500	98.87	100.84	560	381.52	809	277,247.49	40.49	11.01	48.50	4; 7; 12; 23; 28
		1000	99.62	174.80	77	244.52	1048	411,710.46	27.27	7.41	65.32	4; 7; 12; 23; 28
$\lceil \lvert V\rvert/p \rceil$	3	1	98.56	38.42	530	155.02	1073	161,989.74	84.61	15.20	0.18	12; 22; 28
		100	98.54	62.40	289	130.54	675	191,404.41	71.61	12.87	15.52	12; 22; 28
		500	97.94	110.09	871	583.62	1298	309,484.80	45.27	7.18	47.55	11; 22; 28
		1000	97.96	89.70	3015	3447.42	3804	454,614.67	30.74	5.45	63.81	12; 22; 28
	4	1	99.36	25.55	22	34.87	514	145,732.10	81.24	18.54	0.21	12; 23; 26; 28
		100	98.86	33.09	442	117.67	1321	176,241.71	67.46	15.07	17.47	12; 23; 26; 28
		500	98.22	63.24	2782	1521.90	4931	295,787.67	41.19	8.91	49.90	12; 15; 27; 28
		1000	97.95	94.57	10053	4550.67	8508	443,393.74	27.48	5.94	66.58	12; 15; 27; 28
	5	1	98.05	40.70	625	185.50	1115	139,032.42	76.90	22.88	0.22	5; 11; 23; 26; 28
		100	98.20	42.48	3268	630.09	2996	168,736.29	63.39	18.97	17.64	5; 11; 23; 26; 28
		500	98.09	91.67	3340	2472.09	5069	286,728.55	37.70	11.05	51.26	5; 11; 23; 27; 28
		1000	93.69	88.62	9004	**t.l.**	17450	453,254.27	25.01	6.85	68.14	3; 20; 23; 27; 28

Table 7
Results for AP50.

q	p	β	r-gap	r-time	nodes	time	nCuts	totalCost	%-access	%-interHub	%-cycle	hubs
$\|V\|$	3	1	99.75	114.77	17	147.25	993	158,880.67	85.43	14.38	0.20	14; 28; 35
		100	99.59	150.82	94	237.14	1063	189,643.25	71.57	12.05	16.39	14; 28; 35
		500	98.88	395.54	1033	1963.21	2078	313,509.45	43.40	7.38	49.22	14; 28; 35
		1000	99.46	463.62	198	771.47	1536	461,294.25	32.54	4.63	62.83	9; 14; 35
	4	1	99.77	89.87	36	112.20	877	143,692.01	82.38	17.40	0.22	14; 28; 33; 35
		100	99.74	130.99	18	166.33	946	174,356.01	67.96	14.28	17.76	14; 28; 33; 35
		500	99.03	377.09	1685	3351.32	2679	297,364.38	40.01	8.44	51.56	14; 28; 33; 35
		1000	99.02	449.10	666	1822.15	1756	444,031.52	28.45	6.19	65.37	9; 14; 29; 35
	5	1	99.53	89.11	58	119.12	756	132,689.72	77.90	21.86	0.24	4; 14; 28; 33; 35
		100	99.58	153.72	107	222.71	841	163,460.64	63.55	17.47	18.98	4; 14; 28; 33; 35
		500	99.78	292.64	48	369.10	1123	281,644.93	38.61	9.80	51.60	9; 14; 28; 33; 35
		1000	99.40	585.33	337	1152.33	1536	426,543.08	26.06	6.48	67.46	9; 14; 28; 33; 35
$\lceil\|V\|/2\rceil$	3	1	99.74	101.03	14	124.61	919	158,880.67	85.43	14.38	0.20	14; 28; 35
		100	99.62	158.50	34	218.20	970	189,643.25	71.57	12.05	16.39	14; 28; 35
		500	98.87	289.51	808	1644.84	2022	313,509.45	43.40	7.38	49.22	14; 28; 35
		1000	97.96	393.14	2443	**t.l.**	4435	468,204.82	29.00	5.59	65.41	14; 19; 35
	4	1	99.78	99.65	17	126.91	827	143,692.01	82.38	17.40	0.22	14; 28; 33; 35
		100	99.74	142.96	22	182.88	917	174,356.01	67.96	14.28	17.76	14; 28; 33; 35
		500	99.01	355.12	1045	2060.49	1785	297,364.38	40.01	8.44	51.56	14; 28; 33; 35
		1000	98.98	473.82	1230	3966.06	3575	444,031.52	28.45	6.19	65.37	9; 14; 29; 35
	5	1	99.57	91.42	74	133.27	788	132,688.63	77.82	21.94	0.24	4; 14; 28; 33; 35
		100	99.58	144.35	57	192.27	796	163,460.64	63.55	17.47	18.98	4; 14; 28; 33; 35
		500	99.76	322.70	26	409.19	1087	281,644.93	38.61	9.80	51.60	9; 14; 28; 33; 35
		1000	99.33	449.22	428	1106.70	1324	426,543.08	26.06	6.48	67.46	9; 14; 28; 33; 35
$\lceil\|V\|/p\rceil$	3	1	98.98	122.24	483	431.89	1903	162,358.48	84.27	15.53	0.21	14; 28; 35
		100	98.76	263.20	1023	1534.43	1851	193,611.51	70.63	13.16	16.22	14; 28; 35
		500	98.17	234.24	3412	5173.27	6609	318,506.03	43.65	7.53	48.81	14; 27; 35
		1000	95.19	440.67	2315	**t.l.**	6352	487,540.97	30.17	5.43	64.40	13; 29; 35
	4	1	99.69	90.20	77	156.42	729	144,210.66	82.55	17.22	0.22	14; 28; 33; 35
		100	99.44	149.71	503	549.79	1374	175,349.65	67.97	14.12	17.91	14; 28; 33; 35
		500	98.82	242.71	1599	3343.15	7156	300,103.26	40.02	8.15	51.83	14; 28; 33; 35
		1000	96.88	334.48	2090	**t.l.**	6466	462,471.69	27.37	5.11	67.52	14; 33; 35; 39
	5	1	96.88	81.78	6188	2764.23	4354	140,093.97	76.32	23.44	0.24	3; 16; 29; 33; 35
		100	96.20	111.65	12369	**t.l.**	12,712	173,088.85	64.02	17.23	18.76	3; 16; 28; 33; 35
		500	92.65	215.80	6200	**t.l.**	13,972	311,508.49	37.68	8.76	53.56	14; 17; 28; 33; 35
		1000	83.97	257.20	3418	**t.l.**	11,940	524,502.41	24.05	4.86	71.09	14; 17; 28; 34; 36

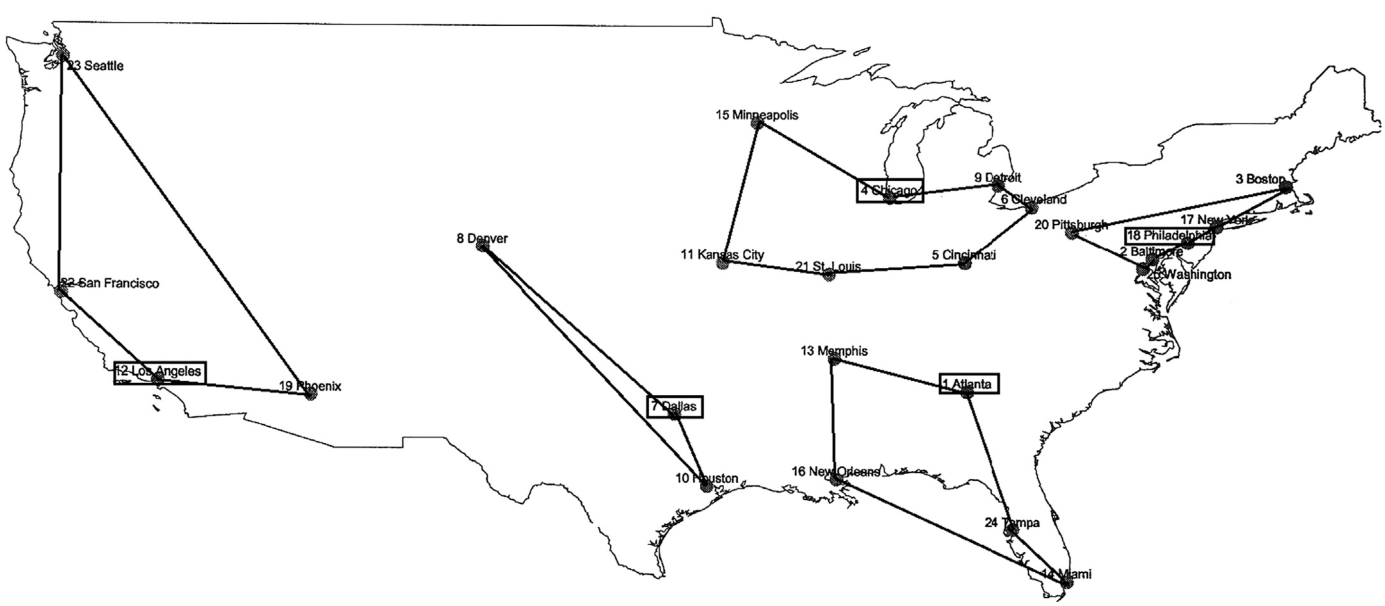

Fig. 2. Optimal solution for CAB25 with $p=5$, $q=13$, $\alpha=0.8$, $\beta=0.01$.

Then, for each pair of nodes $i, i' \in V$ let us define a support graph $G' = (V', E')$ where $V' = V$ and E' contains all edges $(i,j) \in E$ such that $j \in V$ and $x^*_{ij} + 2z^{*2}_{i'j} > 0$, all edges $(i',j) \in E$ such that $j \in V$ and $x^*_{i'j} + 2z^{*2}_{ij} > 0$, and all other edges $e \in E$ such that $x^*_e > 0$. The capacity of the edges in E' is set to the positive value considered for their definition. Let $S \subset V'$ be such that $i \in S$, $i' \notin S$, and $\delta(S)$ is the minimum cut between i and i' in G'. If the capacity of $\delta(S)$ is smaller than $2(\sum_{j \in V} z^{*2}_{ij} + \sum_{j \in V} z^{*2}_{i'j})$, S defines the most violated constraint (17) for i and i'. Therefore, again the separation problem can be solved by performing a max-flow computation for each pair of nodes. The overall complexity of the algorithm is $O(|E||V|^4)$.

3.2.4. *Separation of inequalities (18)*

Inequalities (18) are similar to the rounded capacity inequalities for the capacitated vehicle routing problem [35]. We propose to separate them heuristically as follows. We look for the min-cut set S in each of the connected components of a support graph with node set equal to V and edge set obtained from E by selecting all edges $e \in E$ with $x^*_e > 0$ and giving them those values as capacity. Then we check whether each S, or its complement within the corresponding connected component, gives a violated inequality (18). Moreover, we also check the violation of inequalities (18) for all subsets $S \subset V$ associated with violated constraints (8) and (17).

The bottleneck of this separation procedure is the min-cut computation, which may be applied at most $|V|$ times. Thus, the complexity of this approach is $O(|E||V|^3)$.

3.2.5. *Separation of inequalities (20)*

Inequalities (20) are separated exactly with a procedure that is also based on the classical separation of the subtour elimination constraints for the TSP. We can rewrite (20) as

$$(q-1)x(\delta(S)) + \sum_{i \in S} 2qz^2_{ii} + \sum_{i \notin S} \sum_{j \in V} 2z^2_{ij} \geq 2 \sum_{i \in V} \sum_{j \in V} z^2_{ij}.$$

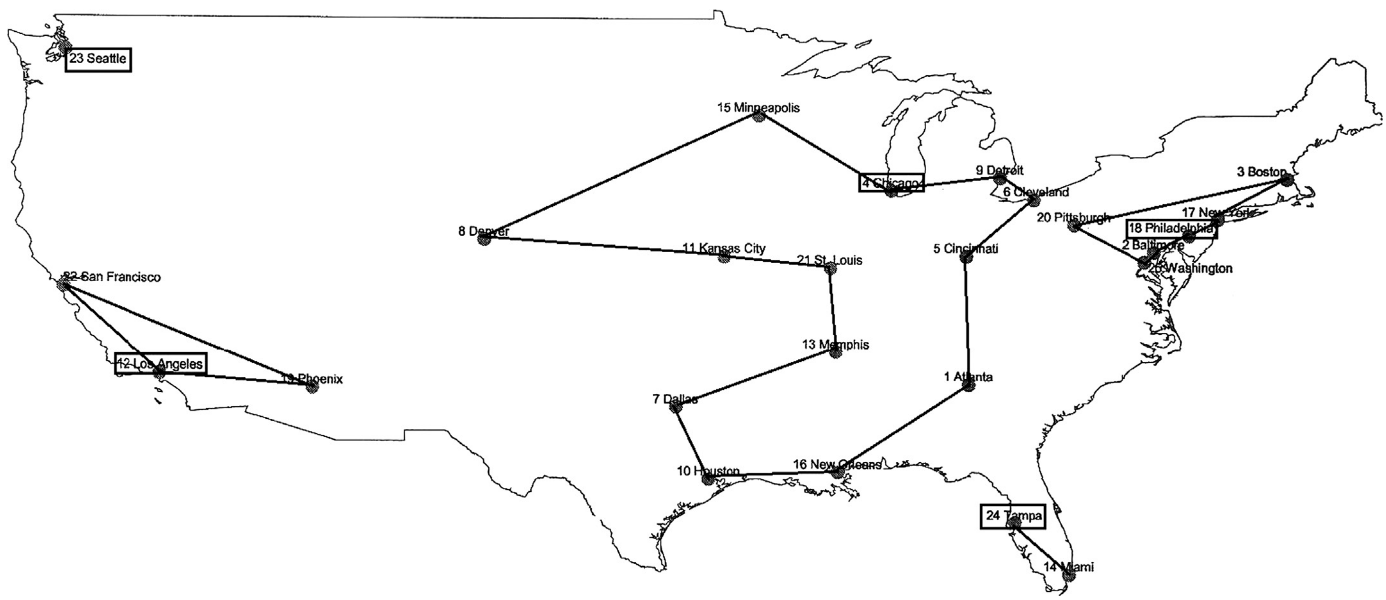

Fig. 3. Optimal solution for CAB25 with $p=5$, $q=13$, $\alpha=0.8$, $\beta=0.05$.

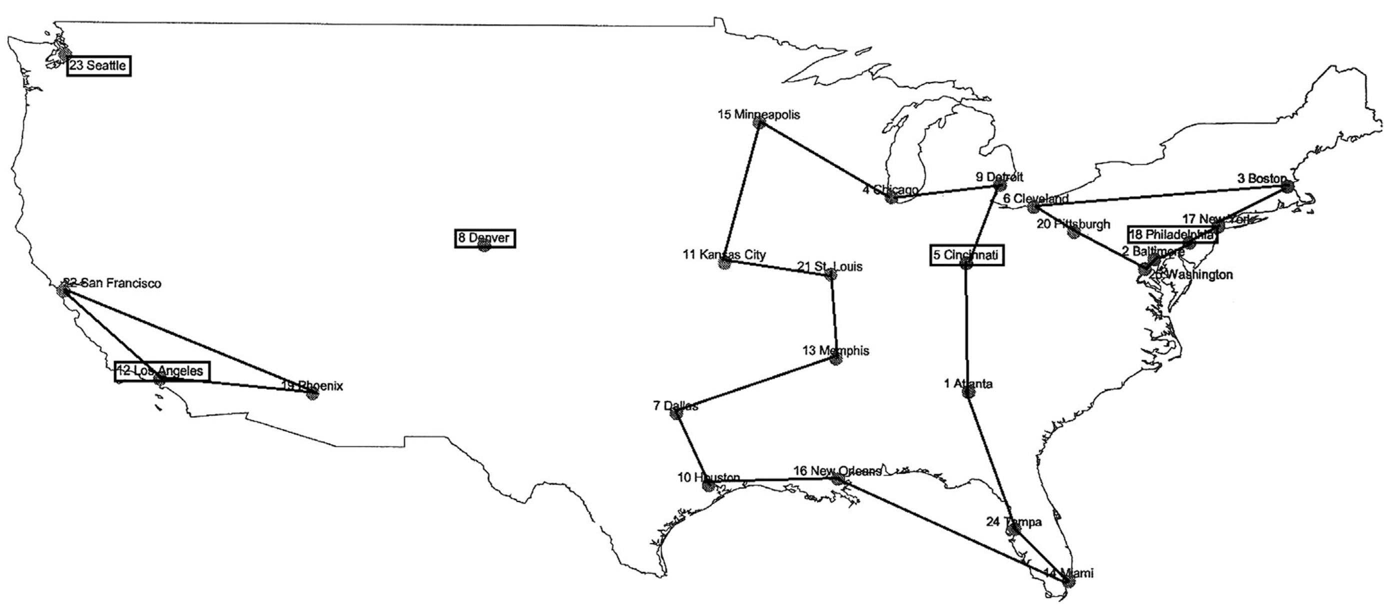

Fig. 4. Optimal solution for CAB25 with $p=5$, $q=13$, $\alpha=0.8$, $\beta=0.2$.

Let us consider a support graph $G' = (V', E')$ with $V' = V \cup \{s, t\}$, s and t being dummy nodes. The edge set E' is obtained by considering

- all edges $e \in E$ such that $x_e^* > 0$, each one with capacity $(q-1)x_e^*$,
- all edges connecting s with nodes $i \in V$, each one with capacity $2qz_{ii}^{*2}$, and
- all edges connecting nodes $i \in V$ with t, each one with capacity $2\sum_{j \in V} z_{ij}^{*2}$.

Then, a set $S \subset V'$ with $t \in S$ and $s \notin S$ defines a violated inequality (20) if the capacity of the cut $\delta(S)$ on G' is smaller than $2\sum_{i \in V}\sum_{j \in V} z_{ij}^{*2}$. Therefore, inequalities (20) can be separated by solving a $s-t$ min-cut problem on G'. The complexity of the separation method is $O(|E||V|^2)$.

4. Computational results

We coded the branch-and-cut algorithm in C++ and ran it on a personal computer with an Intel Core i7 CPU at 3.4 GHz and 16 GB of RAM. We used CPLEX 12.5 as a mixed integer linear programming solver. To solve the min-cut problems we used the implementation of the path-relabel maximum flow algorithm provided by the Concorde TSP solver.

The behavior of the algorithm was first tested on two data sets commonly used in the hub location literature: the US Civil Aeronautics Board (CAB) and the Australian Post (AP) data sets. The CAB data set was introduced by O'Kelly [38] and is based on airline passenger flow among 25 important cities in the US. The AP data set was introduced by Ernst and Krishnamoorthy [22] and is based on postal delivery in 200 postal districts in Sidney,

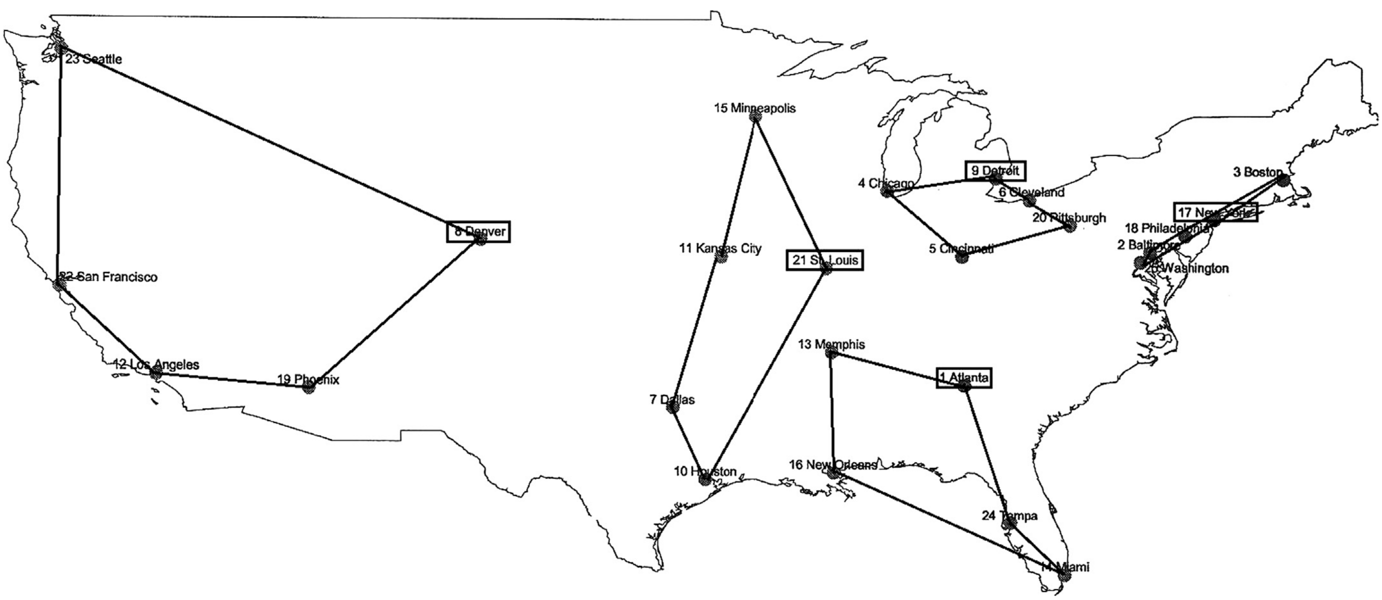

Fig. 5. Optimal solution for CAB25 with $p=5$, $q=5$, $\alpha=0.8$, $\beta=0.2$.

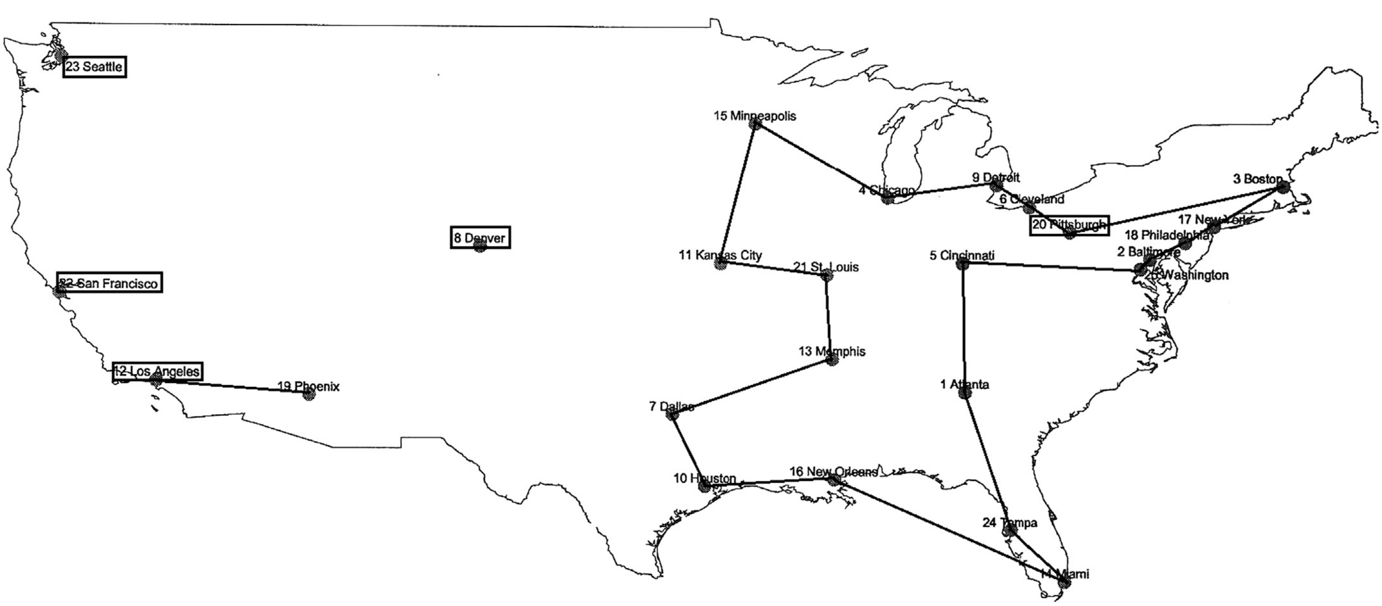

Fig. 6. Optimal solution for CAB25 with $p=5$, $q=25$, $\alpha=0.8$, $\beta=0.2$.

moreover smaller instances can be generated using a code made available by the authors. We generated instances with 25, 40 and 50 nodes. AP data differ from CAB in that the flow matrix of the AP data is not symmetric and flows from a node to itself are positive. Moreover, in AP instances the cost of assigning node i to hub j is set to $3c_{ij}o_i + 2c_{ji}d_i$ and the inter-hub routing cost is discounted with $\alpha = 0.75$.

For all instances we defined c_{ij} and f_{ij} as the Euclidean distance between nodes i and j, and we made experiments with $p = \{3, 4, 5\}$ and $q = \{|V|, \lceil |V|/2 \rceil, \lceil |V|/p \rceil\}$. In CAB instances we set $\alpha = \{0.2, 0.4, 0.8\}$ and $\beta = \{0.01, 0.05, 0.2\}$, while in AP instances $\beta = \{1, 100, 500, 1000\}$. The values of parameter α are the ones commonly used in the literature on hub problems. However, the values of parameter β were chosen in order to obtain optimal solutions with increasing percentage of the circular routes' cost over the total cost and to measure the impact of this on solution time and on the locations of hubs.

Tables 2–7 show the branch-and-cut results. For each instance, columns *r-gap, r-time, nodes, time, nCuts* and *totalCost* report the ratio of the root bound to the optimal value as a percentage, the time spent at the root node, the number of nodes exploited in the branch-and-cut tree, the total solution time in seconds, the total number of cuts added, and the total solution cost, respectively. In the remaining columns we report the assignment, inter-hub routing, and the cycle costs as a percentage of the total cost, as well as the locations of hubs in the best solutions. We imposed a time limit of two hours for each run. When this time limit is exceeded, we report "**t.l.**" in column *time*, we show the cost of the best solution at the end of the computation in column *totalCost* instead of the optimal solution value, and we use that figure to compute the bound in *r-gap*.

Our separation routines found violated inequalities of all families. Inequalities (14)–(16) have simple separation procedures, while (8), (17), (18) and (20) require solving min-cut problems. We

Table 9
SApHMP results for CAB25.

p	α	Access cost	interHub cost	SApHMP	Hubs	Cycles' cost
3	0.2	631.21	136.14	767.35	4; 12; 17	10,233.68
	0.4	637.10	264.60	901.70	4; 12; 18	9813.17
	0.8	657.77	501.07	1158.83	2; 4; 12	9519.25
4	0.2	464.38	165.26	629.63	4; 12; 17; 24	9419.79
	0.4	484.13	303.38	787.52	1; 4; 12; 17	8878.00
	0.8	501.46	586.20	1087.66	1; 4; 12; 18	8878.00
5	0.2	368.18	170.20	538.37	4; 7; 12; 14; 17	8849.93
	0.4	369.89	337.80	707.69	4; 7; 12; 14; 17	8792.14
	0.8	423.23	610.88	1034.10	1; 4; 7; 12; 18	9207.69

Table 8
Comparison between the basic and the complete B&C for AP25.

q	p	β	Basic B&C				Complete B&C			
			r-gap	r-time	nodes	time	r-gap	r-time	nodes	time
$\|V\|$	3	1	99.38	390.74	10	452.92	99.42	3.34	3	3.74
		100	98.68	258.23	74	475.48	99.13	4.17	39	6.30
		500	98.33	286.70	116	642.44	99.54	7.82	18	10.84
		1000	98.44	343.87	202	729.88	99.75	11.14	22	15.91
	4	1	98.90	391.06	59	470.72	98.98	3.15	57	5.74
		100	98.18	280.68	397	859.88	98.73	4.65	129	11.76
		500	98.75	274.90	145	542.06	99.85	6.49	3	8.03
		1000	98.82	242.10	122	591.87	100.00	9.94	0	10.72
	5	1	99.51	194.39	17	229.15	99.52	2.54	12	3.35
		100	99.26	209.21	27	301.19	99.86	4.99	10	6.18
		500	98.99	221.90	62	425.41	99.90	4.68	11	6.13
		1000	99.36	260.77	362	783.14	100.00	3.42	0	3.60
$\lceil \|V\|/2 \rceil$	3	1	99.36	243.55	5	273.64	99.44	3.09	3	3.68
		100	98.63	245.86	77	437.61	99.26	5.76	7	6.52
		500	98.12	269.40	147	701.91	99.31	5.96	13	8.13
		1000	98.34	342.28	476	1050.46	99.55	8.02	36	11.11
	4	1	98.90	383.76	56	482.28	98.94	2.89	35	5.24
		100	97.99	214.70	545	610.59	98.65	3.74	87	7.96
		500	98.81	282.32	342	1021.26	99.92	5.41	6	6.61
		1000	98.86	290.35	80	567.16	100.00	8.13	0	8.28
	5	1	99.54	198.09	23	239.65	99.57	2.53	13	3.48
		100	99.19	220.69	19	325.06	99.93	3.53	5	4.68
		500	98.97	225.22	117	470.19	99.91	4.77	3	5.77
		1000	99.36	228.09	59	415.77	100.00	4.12	0	4.26
$\lceil \|V\|/p \rceil$	3	1	100.00	216.09	0	216.12	100.00	1.84	0	1.86
		100	99.53	256.57	4	333.56	100.00	2.82	0	2.95
		500	97.48	288.88	131	635.00	99.14	11.20	96	18.27
		1000	96.47	287.54	2757	4471.00	98.54	12.29	295	32.76
	4	1	98.82	243.31	6	304.23	99.47	3.14	5	4.12
		100	98.13	204.16	294	597.36	99.50	4.15	5	5.49
		500	96.67	319.01	874	1806.87	98.26	14.15	193	34.71
		1000	95.27	277.68	4451	**t.l.**	97.21	12.07	1335	106.14
	5	1	99.46	57.72	25	112.02	99.68	4.82	12	5.63
		100	97.85	227.00	109	418.38	98.64	3.92	91	7.05
		500	94.71	59.05	6101	**t.l.**	97.40	6.38	757	29.69
		1000	91.31	193.60	5876	**t.l.**	97.52	16.30	2851	166.38

decided to apply the separation procedure for (17) only when no other violated cuts are found due to its time consumption. Still, the violated inequalities (17) found were fundamental to solve some instances to optimality within the time limit.

We observe that the bounds at the root nodes are strong in most cases. All instances except one of AP40 and six of AP50 are solved to optimality within the time limit. The computational experiments show that, for this data, as q gets smaller, the problem becomes more difficult, more nodes are enumerated to reach optimality, and the total cost increases. Also the instances with large β and p are usually more difficult to solve. Indeed, the instance of AP40 that is not solved to optimality within the time limit has tight capacities and large β and p. Out of six unsolved instances of AP50, four have large β value and five have tight capacities.

For the uncapacitated CAB instances, when $\beta = 0.01$ the optimal locations of hubs are very similar to those of the SApHMP. The only differences are for $p=3$ and $\alpha = 0.2$ where the hub at Chicago is moved to Cincinnati, for $p=3$, $\alpha = 0.4$ where the hubs at Chicago and Philadelphia are moved to Cincinnati and New York, and for $p=4$, $\alpha = 0.2$ where the hub at Tampa is moved to Miami, when the cycle costs are included in the objective function. These changes are not over long distances. The empirical test indicates that, in these instances, increasing β has more impact on the locations when α is large.

We look closely into an instance and observe the changes in the location of hubs as β increases. Figs. 2–4 sketch the optimal solutions of the CAB instances with $\alpha = 0.8$, $p=5$, $q=13$ and different values of β. When $\beta = 0.01$, the solution has five cycles with at least three nodes. As β increases to 0.05, Seattle becomes a hub with no other nodes assigned to it. The 13 cities in the interior are covered by one cycle with a hub located at Chicago. The cities in the east, west and south are covered with three smaller cycles. When we further increase β to 0.2, we observe that Denver also becomes a hub with no other node assigned to it and the remaining cities are covered with three cycles. Cleveland joins the cycle in the east and Tampa and Miami join the cycle in the interior. As in this instance the cycle costs f_e satisfy the triangle inequality, covering all nodes by one cycle is a good solution. However, this is not possible due to capacity restrictions. Hence, the solution is covering 13 nodes with one hub leaving further ones as isolated hubs and covering the remaining nodes with smaller cycles. We note that as β increases from 0.01 to 0.2, the contribution of the cycle costs to the total cost increases from 8.18% to 55.26%.

Figs. 4–6 show the effect of the different capacity values on the CAB instance with $\alpha = 0.8$, $\beta = 0.2$ and $p=5$. The case with the smallest percentage of cycle cost corresponds to the uncapacitated instance ($q=25$) depicted in Fig. 6. In that solution there is a large cycle covering 20 cities, a hub with only one node assigned to it,

Table 10
Results for random instances with 25 nodes.

q	p	β	rand25-s1					rand25-s2					rand25-s3				
			r-gap	r-time	nodes	time	nCuts	r-gap	r-time	nodes	time	nCuts	r-gap	r-time	nodes	time	nCuts
$\lvert V\rvert$	3	1	98.72	4.85	31	5.96	325	97.42	5.44	76	12.14	318	98.76	3.95	38	6.57	297
		100	99.31	15.90	29	20.95	566	97.98	8.97	327	29.30	461	99.94	13.49	5	13.87	605
		500	99.96	11.64	4	12.06	563	99.62	17.36	19	19.53	590	99.84	17.92	0	18.25	785
		1000	99.97	18.72	0	18.95	734	99.85	20.73	3	21.31	608	100.00	20.65	0	20.67	885
	4	1	99.84	3.62	8	4.54	245	97.64	7.13	203	20.45	389	99.12	4.88	60	8.44	272
		100	99.77	14.06	4	15.09	521	98.09	9.00	380	37.61	890	99.92	16.33	0	16.49	656
		500	99.97	18.81	2	19.30	609	99.30	14.18	17	17.69	549	100.00	21.76	0	21.84	689
		1000	99.99	27.63	0	27.81	788	99.88	14.96	4	15.52	527	99.87	22.98	0	23.20	913
	5	1	98.93	3.84	62	7.97	210	98.63	7.39	67	13.62	272	99.56	3.78	19	6.02	211
		100	99.40	15.35	11	17.60	516	97.99	10.02	354	26.77	384	99.59	20.84	8	21.90	733
		500	99.16	24.38	26	28.78	729	99.30	15.38	19	19.25	509	98.86	36.46	110	68.84	1056
		1000	98.91	36.64	62	47.05	977	99.04	22.50	48	29.56	653	98.88	50.75	50	60.17	1171
$\lceil \lvert V\rvert/2\rceil$	3	1	98.82	4.63	24	6.43	321	97.26	4.24	76	10.90	285	98.71	4.34	46	7.32	292
		100	99.26	12.34	14	15.02	540	98.26	8.70	149	22.26	460	97.00	23.37	800	152.82	1042
		500	96.88	25.19	1279	217.75	1687	97.83	18.38	219	46.13	717	94.64	63.38	3500	2872.74	3556
		1000	95.72	51.17	3056	868.07	2481	97.21	29.81	1018	210.13	2163	93.42	103.05	4500	**t.l.**	5493
	4	1	99.80	3.74	20	5.13	248	97.54	7.10	286	23.49	412	99.13	6.41	67	9.97	252
		100	99.80	13.68	8	14.82	493	98.18	11.04	394	37.44	661	99.78	24.29	6	25.69	729
		500	99.50	33.01	22	39.33	788	99.00	19.69	69	29.76	643	97.69	49.05	1734	1103.10	3405
		1000	98.45	40.34	166	81.12	1087	99.43	27.50	20	32.93	765	97.16	76.24	1989	1147.42	5182
	5	1	98.82	3.67	100	8.10	207	98.72	7.25	78	13.85	257	99.53	3.57	29	5.57	218
		100	99.40	14.31	24	17.25	543	98.00	9.00	582	36.47	498	99.83	23.09	0	23.49	758
		500	99.72	24.13	7	25.32	759	98.86	19.02	124	29.72	586	98.70	44.91	54	72.57	1168
		1000	99.06	36.02	15	41.54	977	99.34	25.21	46	31.81	674	98.36	73.27	34	113.18	1456
$\lceil \lvert V\rvert/p\rceil$	3	1	97.07	6.66	215	22.45	498	97.28	6.91	81	16.15	370	98.59	5.88	80	14.18	375
		100	92.81	19.70	19,002	3704.74	4058	95.89	14.96	2054	205.87	1918	88.02	42.85	9061	**t.l.**	7688
		500	88.34	34.48	12,241	**t.l.**	10,370	91.86	24.66	23,500	6759.46	5929	76.72	67.72	6787	**t.l.**	10,043
		1000	82.03	42.10	8927	**t.l.**	9407	86.29	34.02	13,312	**t.l.**	9502	83.07	54.04	6224	**t.l.**	9082
	4	1	100.00	2.48	0	2.50	221	97.69	6.29	107	15.55	288	99.49	4.90	25	6.29	232
		100	94.28	24.13	3648	482.06	1946	96.15	11.43	1485	135.03	1568	90.96	36.79	11,119	**t.l.**	7277
		500	89.25	38.14	13,242	**t.l.**	8047	93.05	28.97	8700	2248.77	4842	84.43	46.96	7007	**t.l.**	11,513
		1000	88.56	43.49	9032	**t.l.**	9659	90.81	40.62	11,184	**t.l.**	7413	86.86	62.67	7034	**t.l.**	9601
	5	1	99.01	6.86	272	14.76	608	96.93	4.71	568	23.79	605	98.86	4.32	70	7.27	257
		100	96.13	14.31	2249	183.47	2947	95.37	11.14	1845	76.58	2024	90.27	28.74	14,654	**t.l.**	12,536
		500	89.70	32.04	15,490	**t.l.**	10,644	92.66	20.03	10,644	1379.58	5161	86.78	35.76	9690	**t.l.**	13,176
		1000	86.56	37.50	9274	**t.l.**	11,246	90.02	30.05	19,199	**t.l.**	8978	87.44	41.22	7808	**t.l.**	14,327

and three isolated hubs. On the other extreme, when the capacity is set to the minimum possible value to ensure feasibility (see Fig. 5), the solution is forced to consist of five cycles with five nodes each. Observe that the total computing time needed to solve the instances goes from 9.31 s for $q=25$ to 165.03 s for $q=5$.

To attest the effectiveness of the cuts used in our branch-and-cut scheme, we performed an experiment consisting of comparing a basic version of the algorithm that just solves models (1)–(13), with the default CPLEX settings, to the complete algorithm version. Table 8 shows the performance of the algorithms on AP25 instances. It is clear from the results that the separation of the valid inequalities presented in Section 2 markedly reduces the computation times. In fact, for larger AP instances the basic algorithm is unable to find even a feasible solution within the time limit of two hours.

As mentioned in the introduction, if the costs associated with the cycles are zero, HLRP reduces to SApHMP. Table 9 reports the assignment cost, inter-hub cost, total solution value, and the hub locations, for the known optimal SApHMP solutions of some CAB25 instances. The last column displays the optimal cost of the cycles associated with those solutions. From these data it is possible to calculate the value of a heuristic HLRP solution obtained by solving first the SApHMP and then calculating the optimal cycles. The resulting solution values are worse, as expected, than the optimal HLRP solution values. For example, for $p=3$, $\alpha=0.2$, and $\beta=0.01$, the optimal HLRP value is 858.76, while the heuristic solution value is $767.35+0.01*10,233.68=869.69$; the locations of the hubs also differ in both solutions. This shows the advantage of jointly tackling the hub and routing parts of the problem.

Finally, we observed that in AP and CAB instances the amount of flow originating at each node is highly variable. In fact, in each of the data set CAB25, AP25, AP40 and AP50, there is one node that generates alone as much flow as approximately 40% of the nodes. So, in these instances a few nodes may have a great influence in the hub location decisions.

To see the effect of the flow structure in the problem solution, we generate three random instances as done in Contreras et al. [19]. All instances have 25 nodes with random coordinates in $[0,200]\times[0,200]$, and the costs c_{ij} and f_{ij} are defined as the Euclidean distance between nodes i and j. We consider three types of nodes: low-level (LL) nodes, with total amount of outgoing flow randomly generated in the interval $[0,10]$, medium-level (ML) nodes, with total amount of outgoing flow randomly generated in the interval $[10,100]$, and high-level (HL) nodes, with total amount of outgoing flow randomly generated in the interval $[100,1000]$. The percentages of nodes of each type (LL–ML–HL) in the instances *rand25-s1*, *rand25-s2* and *rand25-s3* are 60%–38%–2%, 35%–35%–30%, and 99%–1%–0%, respectively. We made experiments with $p=\{3,4,5\}$, $q=\{|V|,\lceil |V|/2\rceil,\lceil |V|/p\rceil\}$, the same α and β values used for AP instances, and a time limit of two hours. The results are reported in Table 10. The comments done for CAB and AP instances are still valid for the random instances, but we observe that the problem is harder to solve in *rand25-s3*. As pointed out in [19], a possible explanation is that for those instances there is not a small set of nodes that generates a large amount of flow, and so the decision on the hub locations gets more difficult.

5. Conclusions

In this study we have introduced a variant of the hub location and routing problem, and have proposed an exact solution method. The problem is closely related to the single allocation hub location problem, the plant-cycle location problem and the multi-depot vehicle routing problem, all of which are known to be difficult problems.

We have proposed a branch-and-cut algorithm, which succeeds in solving instances of up to 50 nodes. The emphasis of our research was placed on deriving strong valid inequalities to improve the LP relaxations, and on devising efficient separation procedures. The development of heuristics to tackle larger instances could be an interesting future research direction.

Our experiments have shown that, in our test-bed instances, the problem is more difficult to solve when the number of hubs p to be selected increases. This is coherent with results found in other investigations on related problems, like the classical capacitated vehicle routing problem. Indeed, for these instances a column-generation approach could be a promising alternative to our branch-and-cut algorithm.

Acknowledgments

This research has been partially supported by the research project MTM2012-36163-C06-01. The research of the third author is supported by the Turkish Academy of Sciences.

References

[1] Albareda-Sambola M, Díaz JA, Fernández E. A compact model and tight bounds for a combined location-routing problem. Comput Oper Res 2005;32:407–28.
[2] Alumur S, Kara BY. Network hub location problems: the state of the art. Eur J Oper Res 2008;190:1–21.
[3] Alumur S, Kara BY, Karasan OE. The design of single allocation incomplete hub networks. Transp Res Part B 2009;43:936–51.
[4] Alumur SA, Yaman H, Kara BY. Hierarchical multimodal hub location problem with time-definite deliveries. Transp Res Part E 2012;48:1107–20.
[5] Aykin T. The hub location and routing problem. Eur J Oper Res 1995;83: 200–19.
[6] Baldacci R, Mingozzi A. A unified exact method for solving different classes of vehicle routing problems. Math Program 2009;120:347–80.
[7] Belenguer J-M, Benavent E, Prins C, Prodhon C, Calvo RW. A branch-and-cut method for the capacitated location-routing problem. Comput Oper Res 2011;38:931–41.
[8] Bruns A, Klose A, Stähly P. Restructuring of Swiss parcel delivery services. OR Spektrum 2000;22:285–302.
[9] Calik H, Alumur S, Kara BY, Karasan OE. A tabu-search based heuristic for the hub covering problem over incomplete hub networks. Comput Oper Res 2009;36:3088–96.
[10] Camargo RSd, Miranda Jr. G, Luna HPL. Benders decomposition for hub location problems with economies of scale. Transp Sci 2009;43:86–97.
[11] Camargo RSd, Miranda Gd, Lokketangen A. A new formulation and an exact approach for the many-to-many hub location-routing problem. Appl Math Model 2013;37:7465–80.
[12] Campbell JF. Integer programming formulations of discrete hub location problems. Eur J Oper Res 1994;72:387–405.
[13] Campbell JF, Ernst AT, Krishnamoorthy M. Hub location problems. In: Drezner Z, Hamacher H, editors. Facility location: applications and theory. Berlin: Springer-Verlag; 2002. p. 373–407.
[14] Campbell JF, Ernst AT, Krishnamoorthy M. Hub arc location problems: part I—introduction and results. Manag Sci 2005;51:1540–55.
[15] Campbell JF, Ernst AT, Krishnamoorthy M. Hub arc location problems: part II—formulations and optimal algorithms. Manag Sci 2005;51:1556–71.
[16] Campbell JF, O'Kelly ME. Twenty-five years of hub location research. Transp Sci 2012;46:153–69.
[17] Catanzaro D, Gourdin E, Labbé M, Özsoy FA. A branch-and-cut algorithm for the partitioning-hub location-routing problem. Comput Oper Res 2011;38: 539–49.
[18] Çetiner S, Sepil C, Süral H. Hubbing and routing in postal delivery systems. Ann Oper Res 2010;181(1):109–24.
[19] Contreras I, Cordeau JF, Laporte G. Benders decomposition for large-scale uncapacitated hub location. Oper Res 2011;59:1477–90.
[20] Contreras I, Fernández E, Marín A. The tree of hubs location problem. Eur J Oper Res 2010;202:390–400.
[21] Ebery J. Solving large single allocation p-hub problems with two or three hubs. Eur J Oper Res 2001;128:447–58.
[22] Ernst AT, Krishnamoorthy M. Efficient algorithms for the uncapacitated single allocation p-hub median problem. Locat Sci 1996;4(3):139–54.
[23] Ernst AT, Krishnamoorthy M. An exact solution approach based on shortest paths for p-hub median problems. INFORMS J Comput 1998;10:149–62.
[25] Goldberg A, Tarjan R. A new approach to the maximum-flow problem. J ACM 1988;35:921–40.
[26] Horner MW, O'Kelly ME. Embedding economies of scale concepts for hub network design. J Transp Geogr 2001;9:255–65.

[27] Labbé M, Rodríguez-Martín I, Salazar-González JJ. A branch-and-cut algorithm for the plant-cycle location problem. J Oper Res Soc 2004;55(5):513–20.
[28] Labbé M, Yaman H. Projecting flow variables for hub location problems. Networks 2004;44:84–93.
[29] Labbé M, Yaman H. Solving the hub location problem in a star-star network. Networks 2008;51:19–33.
[30] Labbé M, Yaman H, Gourdin E. A branch and cut algorithm for hub location problems with single assignment. Math Program 2005;102:371–405.
[31] Laporte G, Nobert Y. An exact algorithm for minimizing routing and operating costs in depot location. Eur J Oper Res 1981;6:224–6.
[32] Laporte G, Nobert Y, Arpin D. An exact algorithm for solving a capacitated location-routing problem. Ann Oper Res 1986;6:293–310.
[33] Laporte G, Nobert Y, Arpin D. Optimal solutions to capacitated multidepot vehicle routing problems. Congr Numer 1984;44:283–92.
[34] Laporte G, Nobert Y, Taillefer S. Solving a family of multi-depot vehicle routing and location-routing problems. Transp Sci 1988;22:161–72.
[35] Letchford AN, Eglese RW, Lysgaard J. Multistars, partial multistars and the capacitated vehicle routing problem. Math Program 2002;94:21–40.
[36] Nagy G, Salhi S. The many-to-many location-routing problem. TOP 1998;6:261–75.
[37] Nickel S, Schöbel A, Sonneborn T. Hub location problems in urban traffic networks. In: Pursula M, Niittymäki J, editors. Mathematical methods on optimization in transportation systems. US: Springer; 2001. p. 95–107.
[38] O'Kelly ME. A quadratic integer program for the location of interacting hub facilities. Eur J Oper Res 1987;32:393–404.
[39] O'Kelly ME, Bryan D. Hub location with flow economies of scale. Transp Res Part B 1998;32:605–16.
[40] O'Kelly ME, Miller HJ. The hub network design problem: a review and synthesis. J Transp Geogr 1994;2:31–40.
[41] Podnar H, Skorin-Kapov J, Skorin-Kapov D. Network cost minimization using threshold-based discounting. Eur J Oper Res 2002;137:371–86.
[42] Rieck J, Ehrenberg C, Zimmermann J. Many-to-many location-routing with inter-hub transport and multi-commodity pickup-and-delivery. Eur J Oper Res 2014;236:863–78.
[43] Skorin-Kapov D, Skorin-Kapov J, O'Kelly ME. Tight linear programming relaxations of uncapacitated p-hub median problems. Eur J Oper Res 1996;94:582–93.
[44] Wasner M, Zäpfel G. An integrated multi-depot hub-location vehicle routing model for network planning of parcel service. Int J Prod Econ 2004;90:403–19.
[45] Yaman H. Allocation strategies in hub networks. Eur J Oper Res 2011;211:442–51.
[46] Yaman H. Star p-hub median problem with modular arc capacities. Comput Oper Res 2008;35:3009–19.
[47] Yaman H. The hierarchical hub median problem with single assignment. Transp Res Part B 2009;43:643–58.
[48] Yaman H, Elloumi S. Star p-hub center problem and star p-hub median problem with bounded path lengths. Comput Oper Res 2012;39:2725–32.
[49] Yaman H, Kara BY, Tansel BC. The latest arrival hub location problem for cargo delivery systems with stopovers. Transp Res Part B 2007;41:906–19.
[50] Yoon MG, Current J. The hub location and network design problem with fixed and variable arc costs: formulation and dual-based solution heuristic. J Oper Res Soc 2008;59:80–9.