

Contents lists available at [ScienceDirect](http://ScienceDirect.com)

Computers & Industrial Engineering

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

Comparison of the formulations for a hub-and-spoke network design problem under congestion

Ramez Kian^{a,b,*}, Kamyar Kargar^c^a Department of Industrial Engineering, Istanbul Bilgi University, Istanbul, Turkey^b Southampton Business School, University of Southampton, Southampton, UK^c Department of Industrial Engineering, Bilkent University, Ankara, Turkey

ARTICLE INFO

Article history:

Received 3 April 2015

Received in revised form 28 April 2016

Accepted 19 September 2016

Available online 20 September 2016

Keywords:

Hub-and-spoke networks

Nonlinear congestion cost

Conic quadratic programming

Valid inequalities

ABSTRACT

In this paper, we study the hub location problem with a power-law congestion cost and propose an exact solution approach. We formulate this problem in a conic quadratic form and use a strengthening method which rests on valid inequalities of perspective cuts in mixed integer nonlinear programming. In a numerical study, we compare two well known types of mathematical modeling in the hub-location problems which are solved with different branch and cut strategies. The strength and weakness of the formulations are summarized based on an extensive numerical study over the CAB data set.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

In many-to-many distribution networks, direct connections between every origin-destination pair are not either economical or practical. Therefore, it is often convenient to design networks using a hub-and-spoke structure. To route the traffic in these networks hub points are used, replacing direct connections with fewer indirect connections. Hubs are special facilities acting as consolidation, sorting, switching, and transshipment points in transportation and distribution systems. Instead of direct shipment, flows are concentrated at hub facilities to save transportation cost as a means to achieve economies of scale. Hub location problems cover wide ranges of real cases from the airline traffic flow (Yang, 2009), cargo delivery (Alumur & Kara, 2009), telecommunication and computer networks (Saboury, Ghaffari-Nasab, Barzinpour, & Jabalameli, 2013; Yıldız & Karaslan, 2015) and postal delivery (Ernst & Krishnamoorthy, 1996). Hub location problems deal with location of the hubs and allocation of demand nodes to the located hub facilities. Depending on how non-hub nodes are allocated to hubs, the hub location problems could be divided into two basic types, single allocation and multiple allocation. In multiple

* Corresponding author at: Department of Industrial Engineering, Istanbul Bilgi University, Istanbul, Turkey; Southampton Business School, University of Southampton, Southampton, UK.

E-mail addresses: ramez.kian@enr.bilgi.edu.tr (R. Kian), kamyar.kargar@bilkent.edu.tr (K. Kargar).

allocation networks, each non-hub node is allowed to be allocated to more than one hub, while in the single allocation protocol, each demand node sends and receives flow through a single hub node. Another distinguishing feature is the inclusion of capacities in the model. Hub location problems could be divided into capacitated and uncapacitated problems based on their capacity limitation, which might be imposed on hub nodes (Ebery, Krishnamoorthy, Ernst, & Boland, 2000) or links connecting nodes (Rodriguez-Martin & Salazar-Gonzalez, 2008).

O'Kelly (1986) can be named as the earliest research on hub location problem. Later, O'Kelly (1987) presented the first quadratic formulation for the single allocation p -hub location problem. His model tries to minimize the total transportation cost. Campbell (1994) addressed linear formulations for both single and multiple allocation variants of the problem. Ernst and Krishnamoorthy (1996) formulated the problem in an alternative way with fewer variables and constraints which solves larger problems and shows better computational time performance. For earlier surveys, one can refer to Klincewicz (1998) and Bryan and O'Kelly (1999). The reader is also referred to Farahani, Hekmatfar, Arabani, and Nikbaksh (2013) and Alumur and Kara (2008) for more recent and detailed reviews on the hub location problems. Beside attempts to improve models and provide efficient solution techniques, the hub location field has experienced a high level of activity upon extension of the basic models to rather realistic situations. Issues such as hub location models in competitive context (Gelareh,

Nickel, & Pisinger, 2010), virtual hub location in the case of disruption in hubs (Taghipourian, Mahdavi, Mahdavi-Amiri, & Makui, 2012), multimodal networks (Alumur, Yaman, & Kara, 2012), robust models (Shahabi & Unnikrishnan, 2014), models with fuzzy parameters (Yang, Liu, & Yang, 2013) and congestion on hub networks (De Camargo & Miranda, 2012; De Camargo, Miranda Jr, Ferreira, & Luna, 2009; Elhedhli & Hu, 2005; Elhedhli & Wu, 2010) are among these main extensions. While locating hubs and aggregating flows in order to benefit economic of scale is the main advantage of hub and spoke networks it has an undesired consequence. By consolidating the flow, congestion happens in hub centers or throughout the links. Elhedhli and Hu (2005) study a single hub location problem with congestion cost. They consider a power function as the congestion cost and propose a Lagrangian relaxation based heuristic to decompose the problem into a smaller subproblems. They use an iterative approach to linearize the nonlinear terms in the objective function. Later De Camargo and Miranda (2012) and De Camargo et al. (2009) introduced an adapted congestion cost function in which the congestion cost is imposed only after a given flow threshold. They respectively proposed single and multiple allocation hub location problems with fixed setup costs for hubs and developed a benders decomposition solution approach. In their studies, a Lagrangian approach is employed to solve the subproblems with KKT optimality conditions and get rid of nonlinear terms in the objective function. Then the nonlinear terms are isolated in one of the subproblems and dealt with linearization or KKT and no nonlinear optimization part is fed to the commercial solvers.

All of these works have been modeled as the mixed integer nonlinear programming (MINLP) form based on the formulation of Campbell (1994). This formulation employs more number of variables and constraints than the formulation of Ernst and Krishnamoorthy (1996), which leads to costly computational effort. However, the structure of constraints in Campbell's model is suitable for relaxation and obtaining decomposed sub-problems.

In this paper, we deal with nonlinear terms by employing conic quadratic (MICQP) reformulation instead of linearization method which is used in Elhedhli and Hu (2005). The Campbell based MINLP model is converted to a mixed integer conic quadratic programming which can be efficiently solved via CPLEX and some other solvers. Furthermore, benefiting from the fact that our method does not use any relaxation or decomposition we develop and solve Ernst-Krishnamoorthy based MICQP formulation and demonstrate the dominance of this formulation in larger problems.

We have also benefit from an alternative strengthened version of our conic quadratic formulations which rests on the *perspective cut* as defined in Frangioni and Gentile (2006) and investigated further in Günlük and Linderth (2008). This strengthening method has also been used by Aktürk, Atamtürk, and Gürel (2009) in a scheduling problem, and by Koca, Yaman, and Aktürk (2015) for a stochastic lot sizing problem.

We contribute to the literature with proposing an efficient exact solution method to the problem introduced by Elhedhli and Hu (2005). We propose two mathematical models which are strengthened by applying quadratic reformulation and using perspective cut. Moreover, in our numerical study, we compare both models under two relaxation methods of branch and cut procedure.

The rest of this paper is organized as follows: in the next section we introduce the problem, notations and two different mathematical programming formulations. In Section 3, we discuss how to reformulate the models in the form of conic quadratic programming and how to obtain tighter formulations. In Section 4, we provide an extensive numerical study and sensitivity analysis to compare the efficiency of different formulations and solution approaches. Finally, we summarize our main findings and give directions to future studies.

2. Problem statement and formulations

In this section we provide two mixed integer nonlinear programming formulations for the single allocation p -hub location problem under congestion cost (SApHLC). Given a complete network $G(N, A)$ of $|N|$ nodes and $|A|$ arcs with a deterministic demand on the flow from each node $i \in N$ to $j \in N$, the problem is to locate p numbers of the nodes as hubs and assign each non-hub node to a hub in order to minimize total cost. The flows among the hubs are transited by the discount factor of α . In addition to the basic model, it is assumed that each hub incurs a cost called congestion cost which is a nonlinear convex function of the total flow, u , entering to that hub in the form of $f(u) = au^b$. The distance and demand between node i and j are denoted by C_{ij} and W_{ij} , respectively. In the following formulations, let Z_{ik} be the binary variable which gets 1 when the node i is assigned to the hub k , and takes 0, otherwise. Therefore, total flow arriving to the hub k can be stated as $\sum_{i=1}^N \sum_{j=1}^N W_{ij} Z_{ik}$.

2.1. Campbell's formulation

Let x_{ijkl} denote the assignment of the node i and j to hubs k and l , respectively by taking value of 1. Then our problem can be represented as follows:

$$[\text{SApHLC-C}] \min \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^N \sum_{l=1}^N F_{ijkl} x_{ijkl} + \sum_{k=1}^N a \left(\sum_{i=1}^N \sum_{j=1}^N W_{ij} Z_{ik} \right)^b \quad (1)$$

s. t.

$$\sum_{k=1}^N Z_{ik} = 1 \quad \forall i, \quad (2)$$

$$Z_{ik} \leq Z_{kk} \quad \forall i, k, \quad (3)$$

$$\sum_{k=1}^N Z_{kk} = p, \quad (4)$$

$$\sum_{l=1}^N x_{ijkl} = Z_{ik} \quad \forall i, j, k, \quad (5)$$

$$\sum_{k=1}^N x_{ijkl} = Z_{jl} \quad \forall i, j, l, \quad (6)$$

$$x_{ijkl}, Z_{ik} \in \{0, 1\} \quad \forall i, j, k, l. \quad (7)$$

In the objective function (1), the parameters $F_{ijkl} = C_{ik} + \alpha C_{kl} + C_{ij}$ denote the unit transition cost from node i to j using nodes k and l as their corresponding hubs. The second group of the terms correspond to the total congestion costs in the hub nodes. Constraint (2) ensures that each node should be assigned to exactly one hub; while (3) guarantees that non-hub nodes have to be assigned only to hub nodes. Constraint (4) forces the number of hubs which is a given number p . Eq. (5) indicates that the total flow from i to j reaches to node j from hub l . Similarly (6) ensures that the total flow transited from i to j should first pass through hub k . The last constraint defines the types of the variables. In practice, variables x_{ijkl} do not need to be defined in the discrete form and their LP relaxation leads to an integral solution (see Campbell, 1994).

2.2. Ernst-Krishnamoorthy's formulation

Let Y_{ikl} be the total flow starting from node i and passes through hubs k and l . The parameters O_i and D_i denote the total leaving and entering flow of node i . (i.e., $O_i = \sum_{j=1}^N W_{ij}$, and $D_i = \sum_{j=1}^N W_{ji}$). Then, our problem can also be stated as below:

$$[\text{SApHLC} - \text{EK}] \min \sum_{i=1}^N \sum_{k=1}^N C_{ik} (O_i + D_i) z_{ik} + \sum_{i=1}^N \sum_{k=1}^N \sum_{l=1}^N \alpha C_{kl} Y_{ikl} + \sum_{k=1}^N a \left(\sum_{i=1}^N \sum_{j=1}^N W_{ij} z_{ik} \right)^b \quad (8)$$

s.t.

(2)–(4),

$$\sum_{l=1}^N Y_{ikl} - \sum_{l=1}^N Y_{ilk} + \sum_{j=1}^N W_{ij} z_{jk} = O_i z_{ik} \quad \forall i, k, \quad (9)$$

$$Y_{ikl} \geq 0, z_{ik} \in \{0, 1\} \quad \forall i, k, l. \quad (10)$$

In the objective (8), the first summation group corresponds to the transportation cost between non-hub and hub nodes, the second reflects the inter-hubs costs and the last group, similar to the previous model, indicates the congestion costs. Constraints (2)–(4), appear the same as the previous model with the same purpose. Constraint (9) is called divergence equation of the flow originating from node i at the node k .

3. Conic quadratic reformulation

An optimization problem with a linear objective function and inequality constraints of the form $w^2 \leq uv$ ($u, w \in \mathbb{R}_+$) is called a conic quadratic programming (or, second order cone programming, SOCP) problem. In this section, we describe how we can reformulate our models with general power function terms (such as $U^{\alpha/\beta}$ with $\alpha > \beta, \alpha, \beta \in \mathbb{Q}_+$) in a conic quadratic form. Reformulation of the above models into the SOCP form rests on replacing the nonlinear (power) terms of the congestion costs with the appropriate linear terms in the objective function and adding their equivalent set of cone constraints to the models.

Definition 1.

- a. A cone constraint of the form $w^2 \leq uv$ ($u, w \in \mathbb{R}_+$) is a rotated cone in \mathbb{R}^3 .
- b. A cone constraint of the form $\left\| \begin{matrix} x \\ y \end{matrix} \right\| \leq z$, where the norm $\| \cdot \|$ denotes the Euclidean norm, is a standard cone.

Remark. A rotated cone can be written in the standard form of

$$\left\| \begin{matrix} 2w \\ u - v \end{matrix} \right\| \leq u + v.$$

A function f is conic quadratic representable if its epigraph ($\text{epi}(f) = \{(u, r) | f(u) \leq r\}$) can be equivalently stated by cone constraints. The epigraph inequality of our congestion cost function appears as in (11)

$$U^{a/b} \leq r. \quad (11)$$

Alizadeh and Goldfarb (2003), Nesterov and Nemirovsky (1992) and Nemirovski (2001) provide the related techniques for this purpose to diverse groups of functions. Among those, we know that an inequality of the form

$$t^{2^m} \leq s_1 s_2 \dots s_{2^m} \quad (t \in \mathbb{R}, s_i \in \mathbb{R}_+) \quad (12)$$

can be expressed by at most $2^m - 1$ cone constraints.

The key point is that the exponent of the LHS and the number of variables in the RHS should be equal and a complete power of 2. Thus, the following Lemma is the starting point of our quadratic reformulation procedure. We need to multiply both side of (11) by required number of U variables and then add required number of dummy 1 variables to the RHS.

Lemma 1. An equality of the form $U^{\alpha/\beta} \leq r$ with $\alpha > \beta > 0, r \in \mathbb{R}$ and $U \in \mathbb{R}_+$ is equivalent to the inequality $U^{2^m} \leq r^\beta U^{2^m - \alpha} 1^{\alpha - \beta}$ where $m = \lceil \log_2 \alpha \rceil$.

Proof. Follows directly from algebra.

For illustration, in the following table we provide the conic representation of the power functions that appear in our numerical studies.

To reformulate the [SApHLC-C] and [SApHLC-EK] models in the SOCP form we replace the parenthesis nonlinear congestion cost of hub k , $\left(\sum_{i=1}^N \sum_{j=1}^N W_{ij} z_{ik} \right)^b$, with linear term r_k , add the below auxiliary constraints (13) together with the corresponding quadratic equivalent of the cone constraints as appeared in Table 1 for each hub k .

$$U_k = \sum_{i=1}^N \sum_{j=1}^N W_{ij} z_{ik}. \quad (13)$$

As shown in Table 1, the number of required cone constraints depends on the exponent of the power function which in turn may affect the problem solving time. In other word, dealing with smaller exponents closed to linear function does not mean that we would expect a faster solving time.

3.1. A tighter formulation using perspective cut

Here, we modify the cone representation of the congestion cost based on the following valid inequality

$$U_k^{2^m} \leq r^\beta U_k^{2^m - \alpha} 1^{\alpha - \beta} \leq U_k^{2^m} \leq r^\beta U_k^{2^m - \alpha} z_{kk}^{\alpha - \beta}. \quad (14)$$

When $z_{kk} = 1$, clearly RHS and LHS become the same. While in the case of $z_{kk} = 0$ the equality of LHS and RHS comes from the fact that node k is not a hub and experiences no congestion, that is, $U_k = 0$.

Now, to build cone constraints we start with RHS of (14) and replace all dummy '1' variables in Table 1 with the binary variables z_{kk} . Table 2 depicts this transformation.

This approach provides a tighter formulation and the reason is discussed as follows. The mixed-integer rotated cone constraints include binary variable z , form triples (x, y, z) in $\mathbb{R}^2 \times \mathbb{B}$ and applying LP relaxation would provide their convex hull. This is formally stated in the Lemma 1 of Günlük and Linderoth (2008) which we restate as below proposition.

Proposition 1. The Convex hull of the set $S = \{(x, y, z) \in \mathbb{R}^2 \times \{0, 1\} | x^2 \leq y, Lx \leq x \leq Ux, x, y \geq 0\}$ is $S^c = \{(x, y, z) \in \mathbb{R}^3, x^2 \leq yz, Lx \leq x \leq Ux, 0 \leq z \leq 1, x, y \geq 0\}$ where L and U denote the lower and upper bounds.

Now, we illustrate the two conic quadratic reformulated models for the case of $f(U) = aU^{11/10}$. The formulations for other two

Table 1
Illustration of quadratic reformulation.

Function (f)	$U^{11/10}$	$U^{13/10}$	$U^{3/2}$
Epigraph ($\text{epi}(f)$)	$U^{11/10} \leq r$	$U^{13/10} \leq r$	$U^{3/2} \leq r$
Lemma 1	$U^{16} \leq r^{10} \times U^5 \times 1^1$	$U^{16} \leq r^{13} \times U^3 \times 1^3$	$U^4 \leq r^2 \times U \times 1$
Rotated cones	$U^2 \leq r \times w_1$ $w_1^2 \leq U \times w_2$ $w_2^2 \leq r \times w_3$ $w_3^2 \leq U \times 1$	$U^2 \leq w_1 \times w_2$ $w_1^2 \leq r \times U$ $w_2^2 \leq w_3 \times w_4$ $w_3^2 \leq r \times 1$ $w_4^2 \leq U \times 1$	$U^2 \leq r \times w_1$ $w_1^2 \leq U \times 1$

Table 2

Illustration of tighter quadratic reformulation where $z \in \{0, 1\}$ corresponds to hub opening variable.

Function (f)	$U^{11/10}$	$U^{13/10}$	$U^{3/2}$
Epigraph ($epi(f)$)	$U^{11/10} \leq r$	$U^{13/10} \leq r$	$U^{3/2} \leq r$
RHS of (14)	$U^{16} \leq r^{10} \times U^5 \times z^1$	$U^{16} \leq r^{13} \times U^3 \times z^3$	$U^4 \leq r^2 \times U \times z$
Rotated cones	$U^2 \leq r \times w_1$ $w_1^2 \leq U \times w_2$ $w_2^2 \leq r \times w_3$ $w_3^2 \leq U \times z$	$U^2 \leq w_1 \times w_2$ $w_1^2 \leq r \times U$ $w_2^2 \leq w_3 \times w_4$ $w_3^2 \leq r \times z$ $w_4^2 \leq U \times z$	$U^2 \leq r \times w_1$ $w_1^2 \leq U \times z$

functions (i.e., $f(U) = aU^{13/10}, f(U) = aU^{3/2}$) follow the same logic and are skipped.

$$[C.SApHLC - C] \min \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^N \sum_{l=1}^N F_{ijkl} x_{ijkl} + \sum_{k=1}^N ar_k \quad (15)$$

s.t.

$$(2)-(7),$$

$$(13)$$

$$\left\| \begin{matrix} 2U_k \\ r_k - w_{1k} \end{matrix} \right\| \leq r_k + w_{1k}, \quad \forall k \quad (16)$$

$$\left\| \begin{matrix} 2w_{1k} \\ U_k - w_{2k} \end{matrix} \right\| \leq U_k + w_{2k}, \quad \forall k \quad (17)$$

$$\left\| \begin{matrix} 2w_{2k} \\ r_k - w_{3k} \end{matrix} \right\| \leq r_k + w_{3k}, \quad \forall k \quad (18)$$

$$\left\| \begin{matrix} 2w_{3k} \\ U_k - z_{kk} \end{matrix} \right\| \leq U_k + z_{kk}, \quad \forall k \quad (19)$$

$$U_k, w_{ik} \geq 0, \quad \forall i, k. \quad (20)$$

Table 3

Summary of the formulations performance under different configurations.

	N	r.s.	c.c.	cpu (s)			Epgap (%)			Status (#)			#nodes ave.	
				Min	Max	Ave.o	Min	Max	Ave.f	Opt.	Feasible	Fail		
C.SApHLC-C	10	QCP	n	0.24	313.33	29.70	0.00	0.01	-	24	0	0	775	
		QCP	t	0.21	21.56	3.91	0.00	0.05	-	24	0	0	3	
		LP	n	0.13	3.87	0.66	0.00	0.04	-	24	0	0	69	
		LP	t	0.13	1.38	0.31	0.00	0.05	-	24	0	0	3	
	15	QCP	n	1.93	1196.16	77.42	0.00	48.41	25.03	18	6	0	71	
		QCP	t	1.38	992.08	96.24	0.00	0.04	-	24	0	0	17	
		LP	n	1.35	625.40	45.22	0.00	0.05	-	24	0	0	2104	
	20	LP	t	1.40	18.62	4.25	0.00	0.05	-	24	0	0	15	
		QCP	n	18.07	1196.64	283.06	0.00	62.49	23.79	14	10	0	2	
		QCP	t	8.27	1198.20	237.01	0.00	19.20	9.79	21	3	0	2	
	25	LP	n	8.49	1198.26	87.03	0.00	50.69	35.08	21	3	0	859	
		LP	t	8.95	285.60	44.80	0.00	0.05	-	24	0	0	50	
		QCP	n	68.35	1183.86	82.05	0.00	65.92	48.62	6	5	13	0	
	C.SApHLC-EK	10	QCP	n	0.03	752.22	48.28	0.00	0.04	-	24	0	0	722
			QCP	t	0.05	5.52	1.56	0.00	0.00	-	24	0	0	7
LP			n	0.04	1.51	0.32	0.00	0.03	-	22	0	2	12	
LP			t	0.03	1.09	0.35	0.00	0.04	-	24	0	0	6	
15		QCP	n	0.11	1194.88	45.47	0.00	43.86	23.21	20	4	0	552	
		QCP	t	0.14	206.06	23.61	0.00	0.04	-	24	0	0	37	
		LP	n	0.12	216.25	13.01	0.00	0.05	-	24	0	0	3016	
20		LP	t	0.13	8.19	2.62	0.00	0.05	-	24	0	0	25	
		QCP	n	0.24	1196.50	20.88	0.00	55.22	29.36	18	6	0	173	
		QCP	t	0.21	1175.50	31.40	0.00	5.71	3.75	22	2	0	66	
25		LP	n	0.25	1158.10	5.33	0.00	3.58	3.58	20	1	3	6773	
		LP	t	0.25	29.68	5.82	0.00	0.05	-	24	0	0	52	
		QCP	n	1.00	1196.18	132.37	0.00	60.59	31.32	17	7	0	87	
25		QCP	t	0.95	1176.28	77.38	0.00	7.71	7.64	22	2	0	32	
		LP	n	1.00	1080.82	18.61	0.00	39.03	39.03	17	1	6	2034	
	LP	t	1.10	84.01	19.82	0.00	0.05	-	24	0	0	95		

$$[C.SApHLC - EK] \min \sum_{i=1}^N \sum_{k=1}^N C_{ik}(O_i + D_i)z_{ik} + \sum_{i=1}^N \sum_{l=1}^N \sum_{k=1}^N \alpha C_{kl} Y_{ikl} + \sum_{k=1}^N ar_k \quad (21)$$

s.t.

$$(2)-(4), (9), (10),$$

$$(13),$$

$$(16)-(19)$$

$$(20).$$

In the above formulations, the constraints (16)–(19) are standard cone representation of $U_k^{11/10} \leq r_k$ given in Table 2.

4. Numerical study

The data set used for our computational experiments is the commonly used CAB data set. A single discount factor level $\alpha = 0.2, p \in \{3, 4\}$ number of hubs and the congestion cost function parameters are set to $a \in \{0.1, 1, 10\}$ and $b \in \{1, 1.1, 1.3, 1.5\}$ as in Elhedhli and Hu (2005). The models have been tested for four subsets, $N \in \{10, 15, 20, 25\}$, where each subset includes 24 ($=p \times a \times b = 2 \times 3 \times 4$) instances. The formulations were coded and compiled via C++ language with GCC compiler calling ILOG CPLEX 12.5.1 with ILOG Concert Technology. We have used a PC with 4 GB RAM and Intel(R) Core(TM)2 Quad CPU Q8300 2.5 GHz to execute our codes. The MIQCP barrier optimizer was selected with the time limit of 20 min. All the CPLEX cuts were disabled and the solver's MIP tolerance was set as 0.05%.

The overall results of the Compell [C.SApHLC-C] and Ernst-Krishnamoorthy [C.SApHLC-EK] formulations with different settings are summarized in Table 3. The column labeled with 'r.s.' denotes the relaxation strategy at branch and cut procedure of the solver which is either linear relaxation (LP) or quadratic constraints relaxation (QCP). The column entitled 'c.c.' specifies the

cone constraints types. The normal formulations appears with 'n' and the 't' refers to the tight ones. The 'ave.o' column stands for the average execution time of optimally solved instances. The column 'ave.f' represents the average relative gap percentage between the objective value and best lower bound of the instances which are not optimally solved. As we can observe in this table, the LP-relaxation strategy on branched nodes outperforms the QCP-relaxation approach in terms of running time, number of optimally solved instances and the relative MIP gap of non optimal solutions.

This result holds in both formulations regardless of the cone constraints types. The [C.SApHLC-EK]formulations do not experience any failure due to numerical issues in the QCP solution strategy. On the other hand, [C.SApHLC-C] formulations do not fail to solve with LP relaxation strategy and give a feasible solution in all instances regardless of the quality. From the formulation strength aspect, the tightened models spend less computational effort and also meet less number of branching nodes during the branch and bound procedure (see the last column).

Table 4
Computational results for 10-node network.

N	p	a	b	Cost		cpu		Node		Δ		Optimal hubs
				EK	C	EK	C	EK	C	EK	C	
10	3	0.1	1	492.034	492.034	0.08	0.24	3	0	0.00	0.00	4, 6, 7
10	3	1	1	492.934	492.934	0.06	0.23	3	0	0.00	0.00	4, 6, 7
10	3	10	1	501.934	501.934	0.07	0.24	3	0	0.00	0.00	4, 6, 7
10	3	0.1	1.1	492.274	492.272	1.09	0.14	7	0	0.00	0.00	4, 6, 7
10	3	1	1.1	495.312	495.309	0.71	0.15	1	0	0.00	0.00	4, 6, 7
10	3	10	1.1	525.682	525.68	0.7	0.17	1	0	0.00	0.00	4, 6, 7
10	3	0.1	1.3	495.803	495.8	0.71	0.17	2	0	0.00	0.00	4, 6, 7
10	3	1	1.3	530.593	530.587	0.51	0.4	8	1	0.00	0.00	4, 6, 7
10	3	10	1.3	878.495	878.466	0.53	1.38	40	35	0.02	0.00	4, 6, 7
10	3	0.1	1.5	536.566	536.514	0.32	0.19	7	0	0.06	0.00	4, 6, 7
10	3	1	1.5	933.653	933.652	0.3	0.38	12	3	0.00	0.00	4, 6, 7
10	3	10	1.5	4626.74	4626.64	0.21	0.5	2	0	0.10	0.00	4, 6, 7
10	4	0.1	1	395.23	395.23	0.04	0.24	0	0	0.00	0.00	3, 4, 6, 7
10	4	1	1	396.13	396.13	0.04	0.22	0	0	0.00	0.00	3, 4, 6, 7
10	4	10	1	405.13	405.13	0.03	0.23	0	0	0.00	0.00	3, 4, 6, 7
10	4	0.1	1.1	395.457	395.457	0.16	0.13	0	0	0.00	0.00	3, 4, 6, 7
10	4	1	1.1	398.393	398.392	0.12	0.15	0	0	0.00	0.00	3, 4, 6, 7
10	4	10	1.1	427.75	427.75	0.14	0.16	0	0	0.00	0.00	3, 4, 6, 7
10	4	0.1	1.3	398.616	398.616	0.16	0.15	0	0	0.00	0.00	3, 4, 6, 7
10	4	1	1.3	429.983	429.983	0.24	0.19	0	0	0.00	0.00	3, 4, 6, 7
10	4	10	1.3	743.656	743.656	0.77	0.82	37	22	0.00	0.00	3, 4, 6, 7
10	4	0.1	1.5	432.567	432.561	0.15	0.14	0	0	0.01	0.00	3, 4, 6, 7
10	4	1	1.5	769.444	769.438	0.55	0.31	1	1	0.00	0.00	3, 4, 6, 7
10	4	10	1.5	4049.27	4049.13	0.76	0.59	26	9	-4.36	-4.50	3, 4, 6, 7

Table 5
Computational results for 15-node network.

N	p	a	b	Cost		cpu		Node		Δ		Optimal hubs
				EK	C	EK	C	EK	C	EK	C	
15	3	0.1	1	800.071	800.071	0.24	1.5	7	0	-0.30	-0.30	4, 7, 12
15	3	1	1	800.971	800.971	0.25	1.58	7	0	0.00	0.00	4, 7, 12
15	3	10	1	809.971	809.971	0.26	1.6	7	0	0.00	0.00	4, 7, 12
15	3	0.1	1.1	800.647	800.353	8.19	1.64	20	0	-1.48	-1.78	4, 7, 12
15	3	1	1.1	803.806	803.789	5.11	1.63	9	0	0.02	0.00	4, 7, 12
15	3	10	1.1	838.146	838.146	3.53	1.9	8	0	0.00	0.00	4, 7, 12
15	3	0.1	1.3	805.625	805.613	4.28	1.7	14	0	0.01	0.00	4, 7, 12
15	3	1	1.3	856.403	856.391	5.02	4.33	26	9	0.01	0.00	4, 7, 12
15	3	10	1.3	1335.41	1335.28	8.12	14.87	168	85	0.13	0.00	4, 12, 13
15	3	0.1	1.5	879.884	879.877	1.57	2.28	12	0	0.00	0.00	4, 12, 13
15	3	1	1.5	1540.51	1540.5	1.26	6.08	12	1	0.01	0.00	1, 4, 12
15	3	10	1.5	7224.69	7224.38	1.78	18.62	9	67	0.31	0.00	4, 6, 7
15	4	0.1	1	639.875	639.875	0.14	1.54	0	0	0.00	0.00	4, 7, 12, 14
15	4	1	1	640.775	640.775	0.13	1.4	0	0	0.00	0.00	4, 7, 12, 14
15	4	10	1	649.775	649.775	0.13	1.54	0	0	0.00	0.00	4, 7, 12, 14
15	4	0.1	1.1	640.2	640.156	1.45	1.56	2	0	0.04	0.00	4, 7, 12, 14
15	4	1	1.1	643.602	643.584	3.72	1.56	2	0	0.02	0.00	4, 7, 12, 14
15	4	10	1.1	677.877	677.857	3.29	1.71	2	0	0.02	0.00	4, 7, 12, 14
15	4	0.1	1.3	645.404	645.386	1.3	1.53	2	0	0.01	0.00	4, 7, 12, 14
15	4	1	1.3	695.908	695.882	2.5	3.11	15	1	0.03	0.00	4, 7, 12, 14
15	4	10	1.3	1151.04	1151	5.38	17.7	252	201	0.04	0.00	1, 4, 7, 12
15	4	0.1	1.5	722.723	722.7	2.88	2.08	13	0	0.02	0.00	4, 7, 12, 14
15	4	1	1.5	1286.27	1286.26	1.3	3.95	9	0	0.01	0.00	4, 6, 12, 13
15	4	10	1.5	6311.15	6310.78	1.12	6.69	2	0	-25.38	-25.75	1, 4, 6, 7

For both formulations [C.SApHLC-EK] and [C.SApHLC-C], the objective values and computational times are summarized in Tables 4–7. Also we have compared our objective values with the results obtained by Elhedhli and Hu (2005) in the column Δ . The numbers in this column reflect the difference between our objective values and theirs. The corresponding parameters and optimal hubs in each of the 24 instances are given in these tables. For all subsets, N , in the presence of higher congestion cost ($a = 1$,

$b = 1.5$) a relatively significant improvement are observed with respect to benchmark's objective values. By comparing the execution times in the 'cpu' column it can be concluded that [C.SApHLC-EK] formulation outperforms the other as the size of the network, N , increases.

The execution time comparisons are illustrated in Fig. 1: (a)–(h). The computational effort difference becomes more significant when congestion cost increases. The left part of the figure depicts

Table 6
Computational results for 20-node network.

N	p	a	b	Cost		cpu		Node		Δ		Optimal hubs
				EK	C	EK	C	EK	C	EK	C	
20	3	0.1	1	724.638	724.638	0.26	10.61	0	0	0.00	0.00	4, 12, 17
20	3	1	1	725.538	725.538	0.25	9.9	0	0	0.00	0.00	4, 12, 17
20	3	10	1	734.538	734.538	0.25	9.46	0	0	0.00	0.00	4, 12, 17
20	3	0.1	1.1	724.945	724.945	1.89	11.69	0	0	0.00	0.00	4, 12, 17
20	3	1	1.1	728.608	728.608	1.95	11.88	0	0	0.00	0.00	4, 12, 17
20	3	10	1.1	765.24	765.24	9.13	12.85	1	0	0.00	0.00	4, 12, 17
20	3	0.1	1.3	731.343	731.343	9.83	10.2	1	0	0.00	0.00	4, 12, 17
20	3	1	1.3	792.691	792.59	13.48	18.9	8	2	0.10	0.00	4, 12, 17
20	3	10	1.3	1395.51	1395.27	29.68	233.09	538	590	0.24	0.00	4, 12, 17
20	3	0.1	1.5	839.488	839.488	9.77	16	1	0	0.00	0.00	4, 12, 17
20	3	1	1.5	1782.06	1781.83	5.23	59.19	23	5	-4.65	-4.88	4, 7, 17
20	3	10	1.5	10608.6	10606	11.4	82.73	37	5	-44.82	-47.42	4, 7, 17
20	4	0.1	1	577.721	577.721	0.26	9.87	0	0	0.00	0.00	4, 12, 16, 17
20	4	1	1	578.621	578.621	0.26	8.95	0	0	0.00	0.00	4, 12, 16, 17
20	4	10	1	587.621	587.621	0.25	9.74	0	0	0.00	0.00	4, 12, 16, 17
20	4	0.1	1.1	578.013	578.013	2.28	9.91	0	0	0.00	0.00	4, 12, 16, 17
20	4	1	1.1	581.534	581.534	1.51	10.04	0	0	0.00	0.00	4, 12, 16, 17
20	4	10	1.1	616.749	616.749	1.43	11.08	0	0	0.00	0.00	4, 12, 16, 17
20	4	0.1	1.3	583.65	583.65	1.72	10.84	0	0	0.00	0.00	4, 12, 16, 17
20	4	1	1.3	637.912	637.912	2.46	14.55	3	0	0.00	0.00	4, 12, 16, 17
20	4	10	1.3	1173.41	1173.41	12.27	171.15	409	273	0.00	0.00	4, 12, 16, 17
20	4	0.1	1.5	671.237	671.237	1.84	14.59	2	0	0.00	0.00	4, 12, 16, 17
20	4	1	1.5	1477.45	1477.35	1.44	32.47	0	0	0.10	0.00	4, 12, 16, 17
20	4	10	1.5	9218.93	9216.75	20.77	285.6	220	319	-20.30	-22.48	4, 6, 7, 17

Table 7
Computational results for 25-node network.

N	p	a	b	Cost		cpu		Node		Δ		Optimal hubs
				EK	C	EK	C	EK	C	EK	C	
25	3	0.1	1	767.449	767.449	1.1	69.27	13	0	0.00	0.00	4, 12, 17
25	3	1	1	768.349	768.349	1.2	73.93	13	0	0.00	0.00	4, 12, 17
25	3	10	1	777.349	777.349	1.24	72.36	13	0	0.00	0.00	4, 12, 17
25	3	0.1	1.1	767.769	767.769	8.11	69.56	12	0	0.00	0.00	4, 12, 17
25	3	1	1.1	771.546	771.546	9.41	76.94	11	0	0.00	0.00	4, 12, 17
25	3	10	1.1	809.312	809.312	16.56	81.59	16	0	0.00	0.00	4, 12, 17
25	3	0.1	1.3	774.788	774.788	10.4	53.87	10	0	0.00	0.00	4, 12, 17
25	3	1	1.3	841.97	841.74	29.19	108.38	40	13	0.23	0.00	4, 12, 17
25	3	10	1.3	1501.37	1501.36	84.01	1189.83	738	700	0.01	0.00	4, 12, 17
25	3	0.1	1.5	899.336	899.336	9.59	130	30	0	0.00	0.00	4, 12, 17
25	3	1	1.5	2046.57	2046.4	28.21	517.28	54	61	-2.74	-2.91	4, 12, 18
25	3	10	1.5	12881.9	12880.1	60.97	1187.54	257	197	0.06	-1.74	4, 8, 18
25	4	0.1	1	629.734	629.734	1.16	46.77	5	0	0.00	0.00	4, 12, 17, 24
25	4	1	1	630.634	630.634	1.13	53.12	5	0	0.00	0.00	4, 12, 17, 24
25	4	10	1	639.634	639.634	1.18	57.59	5	0	0.00	0.00	4, 12, 17, 24
25	4	0.1	1.1	630.042	630.04	29.41	52.26	9	0	0.00	0.00	4, 12, 17, 24
25	4	1	1.1	633.7	633.7	12.92	52.16	5	0	0.00	0.00	4, 12, 17, 24
25	4	10	1.1	670.293	670.293	24.14	61.9	9	0	0.00	0.00	4, 12, 17, 24
25	4	0.1	1.3	636.403	636.403	5.72	52.13	7	0	0.00	0.00	4, 12, 17, 24
25	4	1	1.3	697.33	697.33	19.49	121.99	45	2	0.00	0.00	4, 12, 17, 24
25	4	10	1.3	1293.91	1292.17	58.77	1191.29	797	657	1.74	0.00	4, 12, 16, 17
25	4	0.1	1.5	742.974	742.754	10.97	96.36	19	0	0.22	0.00	4, 12, 17, 24
25	4	1	1.5	1705.49	1705.49	21.88	231.89	11	0	0.00	0.00	1, 4, 12, 17
25	4	10	1.5	11108.5	11103.6	28.97	774.88	159	222	-4.82	-9.72	1, 4, 12, 17

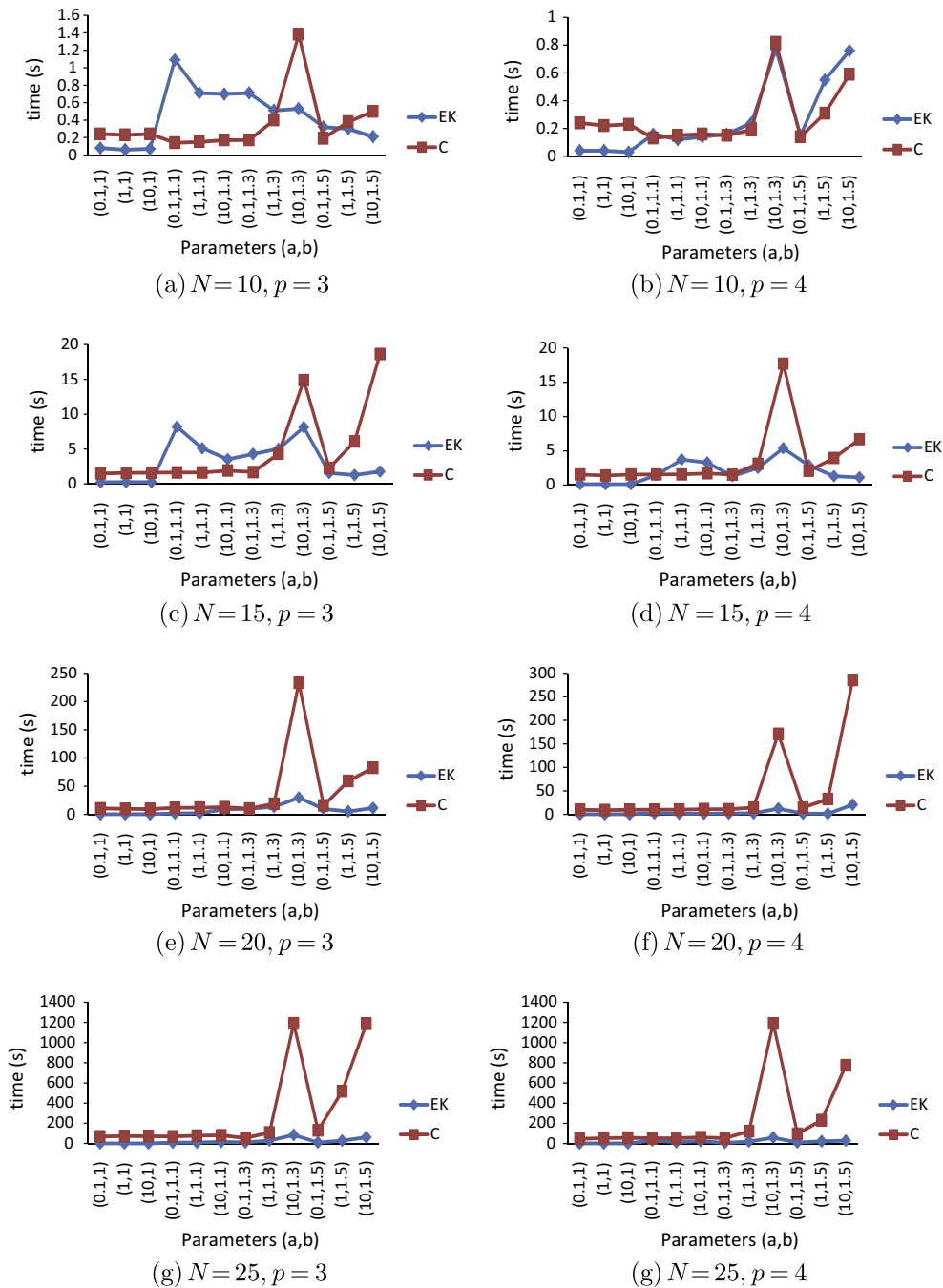


Fig. 1. Execution times of the instances.

the instances with 3 hubs and the right part corresponds to 4-hub instances but both groups follow almost the same trend.

Beside the computational comparisons discussed above, we have investigated the effect of the congestion cost on the flow distribution over the hubs. In the presence of significant congestion costs, traffic distributions become more balanced among the hubs. This pattern is depicted in Fig. 2: (a)–(g) where each bar of the charts presents the proportion of the total traffic passes through each hub.

5. Conclusion and future study directions

In this paper we studied the hub location problem with congestion and we proposed two conic quadratic formulations,

[C.SApHLC-C] and [C.SApHLC-EK]. To the best of our knowledge, it is the first time that Ernst and Krishnamoorthy (1996) based formulation is used for the hub location problem with nonlinear congestion cost. We also made our formulations stronger by applying a type of valid inequality so called perspective cut in mixed integer nonlinear programming. To investigate the behavior of these formulations, two relaxation strategies were imposed. Performances of the proposed formulations were examined over the CAB data set. The numerical results reveal that both of our proposed formulations perform better than the benchmark results of Elhedhli and Wu (2010). Moreover, [C.SApHLC-EK] formulation dominates the [C.SApHLC-C] one in terms of computational effort.

For the future research, the problem can be extended for congestion consideration over connecting links. Also, investigating

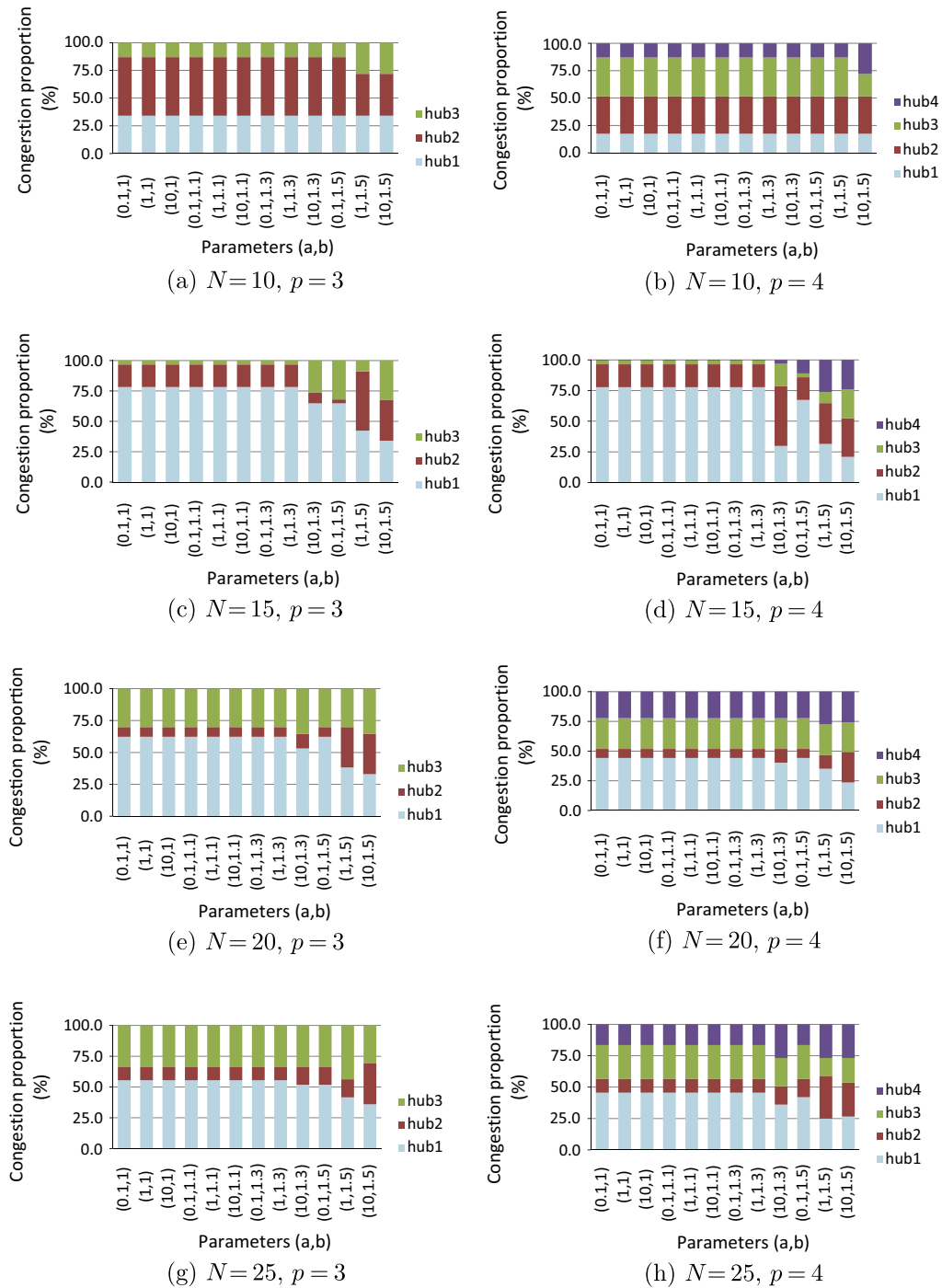


Fig. 2. Flow distribution among hubs.

hub capacity planning models with capacity based congestion cost function would be an interesting research direction.

References

Aktürk, M. S., Atamtürk, A., & Gürel, S. (2009). A strong conic quadratic reformulation for machine-job assignment with controllable processing times. *Operations Research Letters*, 37(3), 187–191.

Alizadeh, F., & Goldfarb, D. (2003). Second-order cone programming. *Mathematical Programming*, 95(1), 3–51.

Alumur, S., & Kara, B. Y. (2008). Network hub location problems: The state of the art. *European Journal of Operational Research*, 190(1), 1–21.

Alumur, S., & Kara, B. Y. (2009). A hub covering network design problem for cargo applications in turkey. *Journal of the Operational Research Society*, 60(10), 1349–1359.

Alumur, S. A., Yaman, H., & Kara, B. Y. (2012). Hierarchical multimodal hub location problem with time-definite deliveries. *Transportation Research Part E: Logistics and Transportation Review*, 48(6), 1107–1120.

Bryan, D. L., & O'Kelly, M. E. (1999). Hub-and-spoke networks in air transportation: An analytical review. *Journal of Regional Science*, 39(2), 275–295.

Campbell, J. F. (1994). Integer programming formulations of discrete hub location problems. *European Journal of Operational Research*, 72(2), 387–405.

De Camargo, R., & Miranda, G. (2012). Single allocation hub location problem under congestion: Network owner and user perspectives. *Expert Systems with Applications*, 39(3), 3385–3391.

De Camargo, R. S., Miranda Jr, G., Ferreira, R. P. M., & Luna, H. (2009). Multiple allocation hub-and-spoke network design under hub congestion. *Computers & Operations Research*, 36(12), 3097–3106.

Ebery, J., Krishnamoorthy, M., Ernst, A., & Boland, N. (2000). The capacitated multiple allocation hub location problem: Formulations and algorithms. *European Journal of Operational Research*, 120(3), 614–631.

- Elhedhli, S., & Hu, F. X. (2005). Hub-and-spoke network design with congestion. *Computers & Operations Research*, 32(6), 1615–1632.
- Elhedhli, S., & Wu, H. (2010). A lagrangean heuristic for hub-and-spoke system design with capacity selection and congestion. *INFORMS Journal on Computing*, 22(2), 282–296.
- Ernst, A. T., & Krishnamoorthy, M. (1996). Efficient algorithms for the uncapacitated single allocation p-hub median problem. *Location Science*, 4(3), 139–154.
- Farahani, R. Z., Hekmatfar, M., Arabani, A. B., & Nikbakhsh, E. (2013). Hub location problems: A review of models, classification, solution techniques, and applications. *Computers & Industrial Engineering*, 64(4), 1096–1109.
- Frangioni, A., & Gentile, C. (2006). Perspective cuts for a class of convex 0–1 mixed integer programs. *Mathematical Programming*, 106(2), 225–236.
- Gelareh, S., Nickel, S., & Pisinger, D. (2010). Liner shipping hub network design in a competitive environment. *Transportation Research Part E: Logistics and Transportation Review*, 46(6), 991–1004.
- Günlük, O., & Linderoth, J. (2008). Perspective relaxation of mixed integer nonlinear programs with indicator variables. In *Integer programming and combinatorial optimization* (pp. 1–16). Springer.
- Klincewicz, J. G. (1998). Hub location in backbone/tributary network design: A review. *Location Science*, 6(1), 307–335.
- Koca, E., Yaman, H., & Aktürk, M. S. (2015). Stochastic lot sizing problem with controllable processing times. *Omega*, 53, 1–10.
- Nemirovski, A. (2001). Lectures on modern convex optimization. In *Society for industrial and applied mathematics*. SIAM. Citeseer.
- Nesterov, Y., & Nemirovsky, A. (1992). Conic formulation of a convex programming problem and duality. *Optimization Methods and Software*, 1(2), 95–115.
- O'Kelly, M. E. (1986). The location of interacting hub facilities. *Transportation Science*, 20(2), 92–106.
- O'Kelly, M. E. (1987). A quadratic integer program for the location of interacting hub facilities. *European Journal of Operational Research*, 32(3), 393–404.
- Rodriguez-Martin, I., & Salazar-Gonzalez, J. J. (2008). Solving a capacitated hub location problem. *European Journal of Operational Research*, 184(2), 468–479.
- Saboury, A., Ghaffari-Nasab, N., Barzinpour, F., & Jabalameli, M. S. (2013). Applying two efficient hybrid heuristics for hub location problem with fully interconnected backbone and access networks. *Computers & Operations Research*, 40(10), 2493–2507.
- Shahabi, M., & Unnikrishnan, A. (2014). Robust hub network design problem. *Transportation Research Part E: Logistics and Transportation Review*, 70, 356–373.
- Taghipourian, F., Mahdavi, I., Mahdavi-Amiri, N., & Makui, A. (2012). A fuzzy programming approach for dynamic virtual hub location problem. *Applied Mathematical Modelling*, 36(7), 3257–3270.
- Yang, K., Liu, Y., & Yang, G. (2013). An improved hybrid particle swarm optimization algorithm for fuzzy p-hub center problem. *Computers & Industrial Engineering*, 64(1), 133–142.
- Yang, T.-H. (2009). Stochastic air freight hub location and flight routes planning. *Applied Mathematical Modelling*, 33(12), 4424–4430.
- Yıldız, B., & Karışan, O. E. (2015). Regenerator location problem and survivable extensions: A hub covering location perspective. *Transportation Research Part B: Methodological*, 71, 32–55.