European Journal of Operational Research 212 (2011) 518–528

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Stochastics and Statistics

Stochastic uncapacitated hub location

Ivan Contreras *, Jean-François Cordeau, Gilbert Laporte

HEC Montréal and Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT), Montreal, Canada H3T 2A7

ARTICLE INFO

Article history:
Received 15 September 2010
Accepted 13 February 2011
Available online 19 February 2011

Keywords:
Hub location
Stochastic programming
Monte-Carlo sampling
Benders decomposition

ABSTRACT

We study stochastic uncapacitated hub location problems in which uncertainty is associated to demands and transportation costs. We show that the stochastic problems with uncertain demands or dependent transportation costs are equivalent to their associated deterministic expected value problem (EVP), in which random variables are replaced by their expectations. In the case of uncertain independent transportation costs, the corresponding stochastic problem is not equivalent to its EVP and specific solution methods need to be developed. We describe a Monte-Carlo simulation-based algorithm that integrates a sample average approximation scheme with a Benders decomposition algorithm to solve problems having stochastic independent transportation costs. Numerical results on a set of instances with up to 50 nodes are reported.

1. Introduction

Hub location problems (HLPs) arise in transportation, telecommunication and computer networks, where hub-and-spoke architectures are frequently used to efficiently route commodities between many origin and destination (O/D) pairs. The performance of these networks relies on the use of consolidation, switching, or transshipment points, called *hub facilities*, where flows from several origins are consolidated and rerouted to their destinations, sometimes via another hub. In HLPs the locations of the hubs as well as the paths for sending the commodities have to be determined. Broadly speaking, HLPs consist in locating hubs on a network so as to minimize the total flow cost.

Due to their multiple applications, these problems are receiving increased attention. Solution methods have been developed for several variants of HLPs analogous to well-known discrete facility location problems, such as uncapacitated hub location, *p*-hub location, *p*-hub center, and hub covering. For each of these classes of problems, there exist several variants arising from various assumptions, such as hub capacities or a specific topological structure for the hub-and-spoke network. There are two basic assumptions underlying most HLPs. The first is that commodities have to be routed via a set of hubs, and thus paths between O/D pairs include at least one hub facility. The second assumption is that hubs are fully interconnected with more effective, higher volume pathways that enable a discount factor τ ($0 < \tau < 1$) to be applied to all transportation costs associated to the commodities routed between a pair of hubs. The reader is refered to Alumur and Kara (2008) and to Campbell et al. (2002) for recent surveys on HLPs.

The location of hub facilities corresponds to long-term strategic decisions which are typically made within an uncertain environment. That is, costs, demands, distances, and other parameters may change after location decisions have been made. Nevertheless, standard HLP models treat data as known and deterministic. This can result in highly sub-optimal solutions given the inherent uncertainty surrounding future conditions. There exist basically two streams of research dealing with optimization under uncertainty: stochastic optimization and robust optimization. In stochastic optimization, it is assumed that the values of the uncertain parameters are governed by known probability distributions. In robust optimization, it is assumed that parameters are uncertain but no information about their probability distributions is know except for the specification of intervals containing the uncertain values.

In classical facility location, stochastic models have been widely investigated over the last four decades. Louveaux (1986, 1993) presents classical reviews on modeling approaches for stochastic facility location in which the location of the facilities is considered as a first-stage decision and the distribution pattern is a second-stage decision. Some of these models (see Louveaux and Peeters, 1992; Laporte et al., 1994) consider capacities on the facilities, and facility size is considered as a first-stage decision. Ravi and Sinha (2006) propose a stochastic problem in which facilities may be open in either the first or second stage, while incurring different installation costs in each stage. The survey by Snyder (2006) covers both stochastic and robust location models for stochastic location problems.

To the best of the authors' knowledge, there exist only three published articles related to stochastic hub location problems. Marianov and Serra (2003) focus on stochasticity at the hub nodes

* Corresponding author.
E-mail addresses: ivan.contreras@cirrelt.ca (I. Contreras), jean-francois.cordeau@cirrelt.ca (J.-F. Cordeau), gilbert.laporte@cirrelt.ca (G. Laporte).

0377-2217/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.ejor.2011.02.018

by representing hub airports as $M/D/c$ queues and limiting through chance constraints the number of airplanes that can queue at an airport. The authors present a linear mixed integer programming (MIP) formulation and propose a heuristic procedure to obtain feasible solutions for instances with up to 30 nodes. Sim et al. (2009) introduce the stochastic p-hub center problem and employ a chance-constrained formulation to model the minimum service-level requirement. Their model takes into account the variability in travel times when designing the hub network so that the maximum travel time through the network is minimized. The authors present a linear MIP formulation for the problem, under the assumption that travel times on the arcs are independent normal random variables. They also propose several heuristics to obtain feasible solutions for instances with up to 25 nodes. Yang (2009) presents a stochastic model for air freight hub location and flight route planning under seasonal demand variations. The author models the problem as a two-stage stochastic program with recourse, in which the location of hub facilities is considered as a first-stage decision and the planning of flight routes as a second-stage decision. Stochasticity on demands as well as on discount factors on the arcs of the network is considered. Moreover, the model allows direct connections between non-hub nodes. A MIP formulation for the problem is presented under the assumption that demand is governed by a discrete probability distribution involving only three possible scenarios. A case study from a 10-node air freight market in Taiwan and China is also described.

One of the problems that have received most attention in deterministic hub location is the *Uncapacitated Hub Location Problem with Multiple Assignments* (UHLPMA). In this problem, the number of required hubs to locate is not known in advance, but a fixed set-up cost for each hub facility is considered. The capacity of the hubs and of the links of the hub network is unbounded. It is assumed that flows originating at the same node but having different destination points can be routed through different sets of hub nodes, i.e. a multiple assignment pattern applies. The objective is to minimize the sum of the hub fixed costs and demand routing costs. The best known formulations for the UHLPMA, in terms of LP relaxation bounds, are those of Hamacher et al. (2004) and of Marín et al. (2006), whereas the best exact algorithms are those of Cánovas et al. (2007), Camargo et al. (2008) and Contreras et al. (2010a). In particular, the Benders decomposition algorithm of Contreras et al. (2010a) can efficiently solve large-scale UHLPMA instances with up to 500 nodes.

In the UHLPMA, demand between O/D pairs as well as transportation costs are treated as known and deterministic. However, in real applications future demand is not known in advance and only a forecast may be available. Transportation costs between node pairs are usually defined to be proportional to the distance between nodes. However, transportation costs are also intimately related to the price of resources (fuel, electricity, raw materials) used to provide the actual transportation of demand, which may be highly uncertain. Other sources of uncertainty in transportation costs may be due to: (i) uncertainty in travel distances, (ii) traffic and congestion, (iii) tariff changes by outsourcing companies, and (iv) link failures. In this paper we study how the UHLPMA can be modeled as a two-stage integer stochastic program with recourse in the presence of uncertainty on demands and transportation costs. In particular, we introduce three different stochastic versions of the UHLPMA. The first is the *Uncapacitated Hub Location Problem with Stochastic Demands* (UHL-SD) in which demands between O/D pairs are considered to be stochastic. The second is the *Uncapacitated Hub Location Problem with Stochastic Dependent Transportation Costs* (UHL-SDC) where uncertainty is given by a single parameter influencing transportation costs. It is assumed that this parameter equally affects the transportation costs for all links of the network. The third is the *Uncapacitated Hub Location Problem with Stochastic Independent Transportation Costs* (UHL-SIC) in which the transportation costs are also stochastic. However, this problem considers the more general case in which the uncertainty of transportation costs is independent for each link of the network. We show that both UHL-SD and UHL-SDC are equivalent to their associated *Expected Value Problem* (EVP) in which uncertain transportation costs are replaced with their expected value (see Birge and Louveaux, 1997). However, this equivalence does not hold for the UHL-SIC.

We use a Monte-Carlo simulation-based method, known as the Sample Average Approximation (SAA) scheme (Kleywegt et al., 2001), to solve UHL-SIC problems with continuous distance distributions, and therefore, an infinite number of scenarios. This method can also be applied to UHL-SIC problems with a finite but very large number of scenarios. We integrate a Benders decomposition scheme to solve the corresponding SAA problems.

The remainder of this paper is organized as follows. Section 2 formally introduces two-stage stochastic models for the considered problems. Section 3 describes our solution method for the UHL-SIC. Computational results are presented in Section 4, followed by conclusions in Section 5.

2. Stochastic uncapacitated hub location problems

Before presenting the stochastic uncapacitated hub location models under study, we describe their deterministic counterpart, the UHLPMA. Let $G = (Q,A)$ be a complete digraph, where Q is the set of nodes and A is the set of arcs. Let also $H \subseteq Q$ represent the set of potential hub locations, and K be the set of commodities whose origin and destination points belong to Q. For each commodity $k \in K$, define W_k as the amount of commodity k to be routed from the origin $o(k) \in Q$ to the destination $d(k) \in Q$. For each node $i \in H$, f_i is the fixed set-up cost for locating a hub at node i. The transportation cost between nodes i and j is defined as $c_{ij} = \gamma d_{ij}$, where d_{ij} is the distance between nodes i and j, which is assumed to satisfy the triangle inequality, and γ is the resource cost per unit distance. All costs relate to the same planning horizon.

Given that hub nodes are fully interconnected and distances satisfy the triangle inequality, every path between an origin and a destination node will contain at least one and at most two hubs. For this reason, paths between two nodes are of the form $(o(k),i,j,d(k))$, where $(i,j) \in H \times H$ is the ordered pair of hubs to which $o(k)$ and $d(k)$ are allocated, respectively. Therefore, the unit transportation cost of routing commodity k along path $(o(k),i,j,d(k))$ is given by $F_{ijk} = \chi c_{o(k)i} + \tau c_{ij} + \delta c_{jd(k)}$, where χ, τ, and δ represent the collection, transfer and distribution costs along the path. To reflect economies of scale between hub nodes, we assume that $\tau < \chi$ and $\tau < \delta$. The UHLPMA consists in locating a set of hubs and in determining the routing of commodities through the hub nodes, with the objective of minimizing the total set-up and transportation cost.

We define binary location variables z_i, $i \in H$, equal to 1 if and only if a hub is located at node i. We also introduce binary routing variables x_{ijk}, $k \in K$ and $(i,j) \in H \times H$, equal to 1 if and only if commodity k transits via a first hub node i and a second hub node j. Following (Hamacher et al., 2004), the UHLPMA can be stated as follows:

$$\text{minimize} \quad \sum_{i\in H} f_i z_i + \sum_{i\in H}\sum_{j\in H}\sum_{k\in K} W_k F_{ijk} x_{ijk} \tag{1}$$

$$\text{subject to} \quad \sum_{i\in H}\sum_{j\in H} x_{ijk} = 1 \quad k \in K \tag{2}$$

$$\sum_{j\in H} x_{ijk} + \sum_{j\in H\setminus\{i\}} x_{jik} \leqslant z_i \quad i \in H, \quad k \in K \tag{3}$$

$$x_{ijk} \geqslant 0 \quad i,j \in H, \quad k \in K \tag{4}$$

$$z \in \mathbb{B}^{|H|}. \tag{5}$$