Computers & Operations Research 37 (2010) 1455–1463

Contents lists available at ScienceDirect

Computers & Operations Research

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

The network p-median problem with discrete probabilistic demand weights

Oded Berman [a], Jiamin Wang [b,*]

[a] Joseph L. Rotman School of Management, University of Toronto, 105 St. George Street, Toronto, Ontario, Canada M5S 3E6
[b] College of Management, Long Island University, C. W. Post Campus, 720 Northern Boulevard, Brookville, NY 11548, USA

ARTICLE INFO

Available online 23 October 2009

Keywords:
Facility location
Network
Median
Probabilistic demand

ABSTRACT

We study the problem of locating p facilities to serve clients residing at the nodes of a network with discrete probabilistic demand weights. The objective is to maximize the probability that the total weighted distance from a client to the closest facility does not exceed a given threshold value. The problem is formulated as an integer program but can be solved only for very small instances because of the exponential number of decision variables and constraints. We analyze the problem and, using a normal approximation of the total weighted distance we develop branch and bound solution procedures for various cases of the problem. We also develop heuristics and meta-heuristics to solve the problem. Based on our computational experiments we make recommendations on which approach to use and when.

1. Introduction

A p-median of a weighted network is defined as a set of p points on the network such that the total weighted distance from the nodal points to a closest point in the set is minimized [11,12,14]. An important application of the p-median model is to optimally locate multiple facilities. In this context, the graph structure describes the underlying transportation network and the nodes, the weight associated with each node and the p-median represent, respectively, the discrete areas where demand originates, the number of potential customers residing in the corresponding area, and the p locations to establish facilities.

Hakimi [12] showed that at least one of the p-medians consisted entirely of nodal points. Kariv and Hakimi [13] suggested a polynomial time algorithm for solving the p-median problem on a tree network. However, finding a p-median on a general network has been proven to be NP-hard [12]. Hence, many publications on the p-median problem are devoted to algorithms and heuristics development. Refer to [15] for an excellent review of some of the typical solution procedures proposed.

Frank [9] argued that the weight associated with each node was not constant in reality. For instance, the minisum criterion is useful for instituting guidelines to locate public libraries. But it is often difficult to precisely estimate the demand (e.g., the number of patrons) coming from each neighborhood. Therefore, it is necessary to treat weights as random variables.

As an extension of the deterministic median concept, Frank [9,10] defined a maximum probability median as a point that maximizes the probability of the total weighted distance being less than or equal to a pre-selected threshold value. When the probabilistic distributions of the demand weights are discrete and independent, Berman and Wang [3] showed that the objective function of the maximum probability median problem is step-wise on each link and hence, the optimum can be found at one of the jump points of the objective function. They then suggested an algorithmic approach to find the optimum by identifying and evaluating dominant jump points.

Later, Berman and Wang [5] studied the maximum probability median problem on a network with demand weights that follow general continuous distributions. They noted that the close form of the objective function was usually difficult to obtain due to convolution operations. Two approximation solution methods—namely normal approximation and discrete approximation were proposed for larger networks.

The maximum probability concept has been applied to extend other classical location models, such as the center and covering problems (see [9,1–4,6,7]) . Frank noted in [9] that even though the definition of a maximum probability median can be readily extended to the p-median case, its solution becomes quite complicated. To the best our knowledge, all the published work on the median and center problems with probabilistic weights is limited to the relatively easier case of $p=1$, i.e., single location. In the present paper, we study the maximum probability p-median ($p>1$) problem when the weights assigned to the nodes are independent discrete random variables.

The remainder of the paper is organized as follows. In the next section, we introduce the maximum probability p-median

* Corresponding author. Tel.: +1 516 299 3914.
E-mail address: jiamin.wang@liu.edu (J. Wang).

0305-0548/$ - see front matter
doi:10.1016/j.cor.2009.10.007

problem and identify conditions under which the optimal solution is simply the deterministic p-median of the network. Since the problem is NP-hard, we develop a normal approximation method, heuristics and relaxation solution procedures in Sections 3–5. Computational experiments are summarized in Section 6 to evaluate the performance of the solution procedures presented in terms of running time and solution quality. Our results indicate that the normal approximation method outperforms other procedures for large-sized instances, while the enumeration procedure and branch and bound algorithm using Lagrangian relaxation tend to solve small-sized instances within a reasonable running time.

2. Problem statement

Let (N, L) be an undirected network, where N is the set of nodes ($|N| = n$) and L is the set of links. Denote by W_h the weight associated with node h $\forall h \in N$. $\{W_h\}$ are assumed to be independent random variables with discrete probability distributions. Suppose that the realizations for any W_h are sorted in ascending order $0 \le w_h[1] < w_h[2] < \cdots < w_h[K_h]$, where K_h is the number of realizations.

It has been noted that nodal optimality does not hold for the maximum probability median problem [9]. For simplicity, we limit the potential sites to the set N. Let C_p be a set of p $(p < n)$ nodes in N. We denote by d_{hj} the shortest distance between any two nodes j and h and by $d(h, C_p) = \min_{j \in C_p}\{d_{hj}\}$ the distance between node h and a closest node in C_p. Let $WD(C_p)$ be the total weighted distance for a given set C_p,

$$WD(C_p) = \sum_{h \in N} W_h d(h, C_p).$$

It is easy to see that $WD(C_p)$ is a discrete random variable. For a given threshold value T, a maximum probability p-median is a set of p nodes, say C_p^*, such that

$$P(WD(C_p^*) \le T) \ge P(WD(C_p) \le T) \quad \forall C_p \subset N.$$

It is evident that C_p^* is the solution to the following model:

$$\min_{C_p \subset N} f(C_p) = P(WD(C_p) > T). \tag{1}$$

The threshold value T can be interpreted as an aspiration level of the total weighted distance traveling from the demand areas to the facilities. The objective of the maximum probability p-median problem is thus to maximize (minimize) the likelihood of achieving (missing) this target.

Let $\mathbf{W} = (W_1, W_2, \ldots, W_n)$ be a vector of independent random weights $W_1, W_2, \ldots, W_n$. Denote by G the set of all possible realizations of $\mathbf{W}$ and by $\mathbf{W}_g$ a vector in G $g = 1, 2, \ldots, |G|$ where the cardinality $|G| = \prod_{h \in N} K_h$ is the number of vectors in G. We further denote by w_{hg} the value of W_h in $\mathbf{W}_g$. Associated with any vector $\mathbf{W}_g$, let

$$P_g = \prod_{h \in N} P(W_h = w_{hg}). \tag{2}$$

Given C_p and $\mathbf{W}_g$, we define

$$Z_g = \begin{cases} 1 & \text{if } \sum_h w_{hg} d(h, C_p) > T, \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

and therefore, the objective function value is calculated as

$$f(C_p) = \sum_{\mathbf{W}_g \in G} P_g Z_g. \tag{4}$$

The propositions presented below show that for certain values of T it is quite easy to find an optimal solution to the problem (1). We note that these results were proved for the special case $p = 1$ in [3] and the proofs for $p > 1$ are very similar and therefore, omitted.

In the first proposition we establish a range of T values for which the deterministic p-median with weight $w_h[K_h]$ (the largest realization of W_h) for any node h, called C_p^0, is optimal for the probabilistic problem.

Proposition 1. *If* $T \ge \sum_{h \in N} d(h, C_p^0) w_h[K_h]$, *then* C_p^0 *is also a maximum probability p-median.*

Let C_p^1 be the deterministic p-median with the weight $w_h[1]$ (the smallest realization of W_h) $\forall h \in N$. In the next proposition we present a range of T values for which any subset of N with p elements is an optimal solution to (1).

Proposition 2. *If* $T < \sum_{h \in N} d(h, C_p^1) w_h[1]$, *any* $C_p \subset N$ *is a maximum probability p-median.*

In the sequel, we will discuss a wide variety of algorithms to find a maximum probability p-median when $\sum_{h \in N} d(h, C_p^1) w_h[1] \le T < \sum_{h \in N} d(h, C_p^0) w_h[K_h]$.

3. Problem formulation and normal approximation

In this section, we formulate the problem (1) as an integer linear program. Exact algorithms such as branch and bound that typically solve the class of integer linear programs become inefficient when the problem is sufficiently large. We thus suggest that the objective function be computed by normal approximation. Revised branch and bound procedures are then developed to solve subproblems derived from normal approximation.

3.1. Formulation

Let X_j be a location variable, where

$$X_j = \begin{cases} 1 & \text{if node } j \text{ belongs to } C_p, \\ 0 & \text{otherwise,} \end{cases}$$

and Y_{hj} be an allocation variable, where

$$Y_{hj} = \begin{cases} 1 & \text{if node } h \text{ is assigned to node } j \in C_p, \\ 0 & \text{otherwise.} \end{cases}$$

The maximum probability p-median problem (1) can be formulated as the following integer linear program (P):

$$\min \sum_{\mathbf{W}_g \in G} P_g Z_g \tag{5}$$

$$\text{s.t.} \sum_{j \in N} X_j = p, \tag{6}$$

$$X_j - Y_{hj} \ge 0 \quad \forall h, j \in N, \tag{7}$$

$$\sum_{j \in N} Y_{hj} = 1 \quad \forall h \in N, \tag{8}$$

$$\sum_{h \in N} w_{hg} \left(\sum_{j \in N} d_{hj} Y_{hj} \right) - M_g Z_g \le T \quad \forall \mathbf{W}_g \in G, \tag{9}$$

$$X_j \in \{0, 1\} \quad \forall j \in N, \tag{10}$$

$$Y_{hj} \in \{0, 1\} \quad \forall h, j \in N, \tag{11}$$

$$Z_g \in \{0, 1\} \quad \forall \mathbf{W}_g \in G. \tag{12}$$