Computers & Operations Research 62 (2015) 237–256

Contents lists available at ScienceDirect

Computers & Operations Research

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

Heuristics for a continuous multi-facility location problem with demand regions

Derya Dinler, Mustafa Kemal Tural, Cem Iyigun

Department of Industrial Engineering, Middle East Technical University, Ankara, Turkey

ARTICLE INFO

Available online 18 September 2014

Keywords:
Facility location problems
Second order cone programming
Minimum sum of squares clustering
Hyperbolic smoothing

ABSTRACT

We consider a continuous multi-facility location allocation problem where the demanding entities are regions in the plane instead of points. The problem can be stated as follows: given m (closed, convex) polygonal demand regions in the plane, find the locations of q facilities and allocate each region to exactly one facility so as to minimize a weighted sum of squares of the maximum Euclidean distances between the demand regions and the facilities they are assigned to.

We propose mathematical programming formulations of the single and multiple facility versions of the problem considered. The single facility location problem is formulated as a second order cone programming (SOCP) problem, and hence is solvable in polynomial time. The multiple facility location problem is NP-hard in general and can be formulated as a mixed integer SOCP problem. This formulation is weak and does not even solve medium-size instances. To solve larger instances of the problem we propose three heuristics. When all the demand regions are rectangular regions with their sides parallel to the standard coordinate axes, a faster special heuristic is developed. We compare our heuristics in terms of both solution quality and computational time.

1. Introduction

The *single-facility location problem* (SFLP) consists of determining the location of a single facility that will serve a set of customers (demand points or regions) while minimizing some objective function, usually a function of the distances between the facility and the locations of the customers. The Fermat–Weber problem or *Weber problem* is a well known SFLP which concerns locating a facility with an objective of minimizing the weighted sum of the (Euclidean) distances between the facility and the customers (weights are positive scalars), see [1,2]. A common approach for solving the problem is known as the *Weiszfeld method* [3]. It is an iterative method that expresses and updates the facility location as a convex combination of the locations of the customers.

In the *multi-facility location allocation problem* (MFLP) each customer is assigned to a single facility and the problem is to determine the optimal locations of a (given) number of facilities, as well as the optimal assignments of customers to facilities.

In some MFLP cases the facilities can be located at a given subset of the plane, in particular at a finite list of candidate sites. This constrained, or *discrete*, MFLP was first considered by Hakimi [4], and is often solved by mixed integer programming, see [5]. In the *continuous*, or unconstrained, MFLP, however, facilities can be located anywhere in the plane.

The continuous MFLP can also be considered as a *clustering problem*, where customers and facility locations correspond to data points and centers, respectively, and a clustering criterion is to be minimized. If the clustering criterion is to minimize the sum of squared distances then MFLP reduces to the *minimum sum of squares clustering problem*.

We study here a continuous multi-facility location allocation problem, where demanding entities are *demand regions* (instead of demand points) in the plane. We assume that each facility has unlimited capacity and can handle all the customers assigned to it. For the following three cases, it would be more appropriate to represent a demanding entity as a region instead of a fixed point:

1. The size of the demanding entity may not be negligible with respect to the distances in the problem.
2. The location of the demanding entity may follow a bivariate distribution on the plane.
3. The number of demanding entities may be so large that it may be more appropriate first to cluster them into regions instead of treating each one separately.

The problem we consider can be stated as follows: given m demand regions in the plane and a positive weight for each region, find the locations of $q \geq 1$ facilities and allocate each region to

E-mail addresses: dinler@metu.edu.tr (D. Dinler), tural@metu.edu.tr (M.K. Tural), iyigun@metu.edu.tr (C. Iyigun).

http://dx.doi.org/10.1016/j.cor.2014.09.001
0305-0548/

exactly one facility so as to minimize the weighted sum of squares of the maximum Euclidean distances between the demand regions and the facilities they are assigned to.

The *maximum distanced point* of a region to a facility does not change if the convex hull of the demand region is taken. Therefore, we restrict ourselves to the case where the demand regions are *convex regions*. Since any region can be approximated in any accuracy by a polygon, here we represent each demand region as a (*closed*) *convex polygon*.

This problem (with polygonal demand regions) is NP-hard in general. When each demand region is a single point, the problem reduces to the minimum sum of squares clustering problem which is NP-hard, in general [6].

We first show that the single facility location problem (with polygonal demand regions) can be solved in polynomial time as it can be modeled as a *second order cone programming* (SOCP) problem. We then formulate the multiple facility location problem as a *mixed integer SOCP problem*. The formulation is a big-M formulation and is weak. It does not even solve medium-size instances within 5 h. To be able to find good solutions for larger problem instances, we propose three heuristics. For small size instances, we compare the solutions of the heuristics with the exact solutions. Moreover, a special case heuristic is developed when the regions are of rectangular shape with sides parallel to the standard coordinate axes.

First heuristic is an *alternate location–allocation* heuristic. We call it as the *SOCP based alternate location allocation heuristic* (SOCP-H). When the locations of the facilities are given, each region is assigned to a facility that minimizes the maximum distance between the region and the facility (*allocation step*). When the allocations of the regions to the facilities are known, the location algorithm solves q SOCP problems to determine the location of the facilities (*location step*). Starting with an initial placement of the facilities, this heuristic repeats allocation and location steps until a stopping condition is reached.

Second heuristic also follows alternate location–allocation scheme. It is named as the *max point based alternate location allocation heuristic* (MP-H) and has the same allocation step with SOCP-H. It is different from SOCP-H in the location step where only one point from each demand region is taken into consideration. When the allocations are known, updated location of a facility is computed by averaging the farthest points of the allocated regions from the previous location of the facility (*location step*). Again, allocation and location steps alternate as in SOCP-H.

The mathematical modeling of the problem is nondifferentiable. Third heuristic called as the *smoothing based heuristic* (SBH) is based on a smoothing strategy which substitutes nondifferentiable functions with continuously differentiable functions. We convert the smoothed problem into an unconstrained nonlinear problem using the implicit function theorem and then solve it with a quasi-Newton algorithm.

The last heuristic is proposed for the special case of *rectangular regions* that is also an alternate location allocation heuristic and is entitled as the *line based heuristic* (LBH). It has the same allocation step with the previous alternate location allocation heuristics. But its location step is quite different. When the allocations of the regions to the facilities are known, algorithm solves q single-facility location problems by converting each one into two single facility location problems on the line. The optimal solutions of these two problems on the line give the coordinates of the optimal solution of the single facility location problem on the plane.

The plan of the paper is as follows: Section 2 reviews the previous work on SFLP and MFLP with demand regions. Section 3 introduces the notation and describes the multi-facility location problem studied in this paper. In Section 4, mathematical programming formulations of the single and multiple facility location problems are presented. Section 5 describes the details of the proposed heuristics. Section 6 introduces a special heuristic for the case with the rectangular demand regions. Section 7 presents the results of the computational experiments. Finally, Section 8 concludes and provides future research directions.

2. Literature review and related work

The distance between a region and a facility can be measured in various ways. Commonly one of the following three different distance definitions is used for measuring the distance between a demand region and a facility, see [7]:

- maximum (farthest) distance,
- minimum (closest) distance,
- expected distance.

In the literature, *expected distance* has been extensively used in solving MFLP with demand regions. This distance may be meaningful when the distance from each facility to every point in each region is important.

Love considered the situation in which the number of demand points is too large to treat each one as a discrete point [8]. The author introduced the possibility of grouping the demand points into demand areas and divided the total population area under consideration into rectangular regions with known dimensions. The objective is to find the location of a facility so as to minimize total expected Euclidean distances between the rectangular regions and the facility. The author proved that the objective function is convex and developed a response-surface technique for solving the problem.

In [9], the authors extended the study in [8]. Love assumed that each demand region is rectangular and has uniform population density. These assumptions ease the complex integral expressions. However in [9], these complex expressions were handled by replacing the demand regions with their centroids. These demand regions are not necessarily rectangular and the population need not be uniformly distributed over the demand regions.

Cooper [10] considered a stochastic extension of the Weber problem. In his paper, the location of the demanding entities was not predetermined but random variables with a given probability distribution and the problem is to minimize the sum of the expected values of the Euclidean distances between the demanding entities and the facility. The probability distribution used in the study was a bivariate normal distribution with two uncorrelated random variables.

Aly and Marucheck [11] tried to locate one or more facilities to serve existing rectangular regions using rectilinear norm. The objective is to minimize total weighted expected distances. The problem was decomposed into two subproblems, one for each coordinate. The objective function was shown to be convex and nondifferentiable and a gradient-free direct search method was proposed for solving the problem.

Carrizosa et al. [12] introduced a general notation for a class of problems where both demanding entities and the facilities can be regions. The notation was inspired by Kendall's notation in Queuing Theory. The objective function considered was to minimize the sum of the expected distances. The work showed the similarities and the differences between the generalized Weber problem and its classical point version.

In the regional Weber problems, evaluation of the objective function, i.e., calculation of the bidimensional integral, has high computational cost. To avoid this, approximation by centroids or disks centered at centroid was used in the literature. Former approximation is mostly dependent to the norm used and shape