European Journal of Operational Research 000 (2014) 1–9

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Discrete Optimization

On the unified dispersion problem: Efficient formulations and exact algorithms

Ting L. Lei*, Richard L. Church

Department of Geography, University of California, Santa Barbara, Santa Barbara, CA 93106-4060, United States

ARTICLE INFO

Article history:
Received 24 June 2013
Accepted 12 October 2014
Available online xxx

Keywords:
Location
Facility dispersion

ABSTRACT

Facility dispersion problems involve placing a number of facilities as far apart from each other as possible. Four different criteria of facility dispersal have been proposed in the literature (Erkut & Neuman, 1991). Despite their formal differences, these four classic dispersion objectives can be expressed in a unified model called the partial-sum dispersion model (Lei & Church, 2013). In this paper, we focus on the unweighted partial sum dispersion problem and introduce an efficient formulation for this generalized dispersion problem based on a construct by Ogryczak and Tamir (2003). We also present a fast branch-and-bound based exact algorithm.

1. Introduction

Facility dispersion problems involve maximizing the separation between facilities to minimize the negative impact they have on each other, minimize potential interaction among hazardous facilities, or enhance the reliability of service and logistic systems. Facility dispersal can be used in a variety of applications including military defense, franchise location, transportation of hazardous materials, layout planning for explosive chemicals (Curtin & Church, 2006) and telecommunication network design (Kim, 2012). Curtin and Church (2007) demonstrate that patterns in the classic central place theory can be replicated using facility dispersal. Facility dispersal has also been deemed as a means to promote the robustness and reliability of critical facilities when it is integrated with other classic location models (Maliszewski, Kuby, & Horner, 2012).

To capture the dispersive quality in different applications, multiple models for facility dispersal have been proposed. Four basic constructs have been developed in the literature to disperse facilities from each other. The first was suggested by Shier (1977) in which *p*-facilities are located while maximizing the minimum distance separating any two facilities. This problem has been called the *p*-dispersion problem (Moon & Chaudhry, 1984) and has also been termed the Max–Min–Min problem (Erkut & Neuman, 1991). Moon and Chaudhry (1984) proposed a second form of facility dispersal, which involved defining the minimum separation distance for each located facility. They proposed to locate *p*-facilities in order to maximize the sum of these minimum separation distances. Moon and Chaudhry called this the *p*-defense problem and Erkut and Neuman classified this problem as a Max–Sum–Min, as this involves MAXimizing the SUM of MINimum separation distances. Kuby (1987) introduced a third form of *p*-facility dispersal that he called *p*-dispersion sum. This problem involved maximizing the sum of all separation distances, which has been called the Max–Sum–Sum problem as it involves MAXimizing the SUM (over each facility) of SUMs (the sum of all separations distances for that facility) (Erkut & Neuman, 1991). The last of the four classic facility location dispersal problems was suggested by Erkut and Neuman (1991). This problem deals with the location of *p*-facilities while maximizing the smallest of the facility defined sums (a facility sum represents the sum of separation distances from a specific facility location to all other facilities) and has been classified as the Max–Min–Sum problem (Erkut & Neuman, 1991).

Curtin and Church (2006) added a new dimension to facility dispersal by conceptualizing a multi-type dispersion metric recognizing that the extent to which two facilities ought to be dispersed should depend on their types. The strength of interaction between facilities is modeled using both the traditional separation distance and a type-specific "repulsion" factor in which a smaller repulsion measure reflects a stronger inter-facility interaction. Curtin and Church developed four dispersion models based on the four basic dispersion metrics classified in Erkut and Neuman (1991) and the multi-type metric where each multi-type dispersion model is a multi-type extension of a classic dispersion model.

Fernández, Kalcsics, and Nickel (2013) proposed an extended dispersion problem involving locating multiple groups of facilities and applied the multi-group dispersion model to the location of recycling facilities. A dispersion metric corresponding to the Max–Min–Min criterion is maximized for each group of facilities. To ensure even distribution of workload, each facility is assigned a weight value and the sum of weights for each group is constrained to be within a range of

* Corresponding author. Tel.: +1 8058934519.
E-mail addresses: tinglei@geog.ucsb.edu (T.L. Lei), church@geog.ucsb.edu (R.L. Church).

http://dx.doi.org/10.1016/j.ejor.2014.10.020
0377-2217/

a pre-specified target weight value. Equity issues in dispersion modeling have also been addressed in Prokopyev, Kong, and Martinez-Torres (2009).

Dispersion problems are related to the maximum independent set problem and the maximum clique problem in graph theory. The p-dispersion problem, for example, can be reduced to the maximum independent set problem by removing edges longer than a given length r and find the maximum independent set on the converted graph (Erkut, 1990). The p-dispersion problem can be solved by solving a series of maximum independent set problems with increasing r values until the size of the maximum independent set is p. Solution methods for the maximum independent set problem have been explored by many researchers. Feo, Resende, and Smith (1994) proposed a parallel Greedy Randomized Adaptive Search Procedure (GRASP) for the maximum independent set problem and found it to be superior in performance to tabu search and simulated annealing methods. Hifi (1997) developed a genetic algorithm for the weighted maximum independent set program, which can be used also to solve equivalent problems such as the maximum clique problem. Gamarnik and Goldberg (2010) presented complexity results of greedy randomized algorithms on constant degree regular graphs.

With few exceptions, the majority of research on facility dispersion in the past two decades has involved the four basic dispersion models classified by Erkut and Neuman (1991). Lei and Church (2013) recently proposed a generalized dispersion model based on the concept of partial sums. At the individual facility level, the partial-sum dispersion metric involves accounting for the sum of the smallest L distances to neighboring facilities, weighted by a set of propulsion factors that can be used, for example, to emphasize the interaction of closely located facilities. At the system level, the model considers the sum of the K smallest facility-defined partial sums. This general construct is called the MaxPSumPSum problem (where PSum stands for partial sum). Lei and Church provided an integer linear programming formulation of this general partial-sum dispersion problem based on assignment variables which tracked specific separation distances.

The partial sum dispersion metric is a compromise of the one-or-all approach found in the four basic dispersion models and makes more sense in modeling many types of inter-facility interactions. For example, in franchise branch location, stores from the same franchise chain should be located away from each other to minimize cannibalization within the same organization. However, existing dispersion models either considers competition from the nearest store and ignores competition from other nearby stores, or considers competition from all stores in the franchise chain including remote sites that have little effect on a store. It would make more sense to consider a number of closest stores using the partial-sum dispersion metric. As another example, in military defense, it is common wisdom to locate assets such as missile silos apart from one another to minimize the chance that two facilities are destroyed at the same time. However, if avoiding the simultaneous loss of two facilities leads to low average spacing between sites, from management's perspective, it may be desirable to devise a facility layout that minimizes the chance of losing three or more facilities simultaneously.

From a theoretical point of view, the MaxPSumPSum construct unifies all four classic dispersion models as special case problems. This means that the MaxPSumPSum problem not only defines a family of new (PSum) dispersion models, but also makes it possible to solve all four existing facility dispersal problems as special case instances. It should be noted that partial-sum dispersion should not be confused with similar partial-sum metrics in median location problems, which involve ***minimizing*** partial sums of demand-to-facility distances either at the individual facility level (Weaver & Church, 1985) or at the system level (Nickel & Puerto, 1999). The reader is referred to Lei and Church (2014) for a median location problem that considers such location criteria in a unified construct.

Facility dispersion problems often have high computational complexities. It is well-known that both the p-dispersion problem and the p-dispersion sum problem are NP-hard (see e.g. Erkut, 1990; Pisinger, 2006). In addition, no polynomial-time heuristic procedure (or approximation algorithm) can guarantee to obtain a near optimal solution for the p-dispersion problem that is within any fixed percentage of the optimal value (Tamir, 1991). In the special case where the separation distances satisfy the triangle inequality, a heuristic for the p-dispersion problem exists with an approximation ratio of 2 (Tamir, 1991) and this ratio is the best possible (Ravi, Rosenkrantz, & Tayi, 1994). An approximation algorithm for the p-dispersion sum problem also exists with an approximation ratio of 4 (Ravi et al., 1994).

Since it subsumes the four classic dispersion problems as special-cases, the generalized dispersion problem should have a computational complexity that is no less than the special-case problems. Lei and Church (2013) report high computation costs of the generalized dispersion problem in their experiments. Certain medium-sized instances of the MaxPSumPSum problem cannot be solved using their ILP formulation in hours. In fact, one example presented in Lei and Church (2013) involving 55 candidate sites and locating 10 dispersive facilities took a week and half to solve. In many applications, such high computational costs may well prevent the model from being applied in practical analysis.

This article aims at operationalizing the generalized dispersion model by developing improved model formulations and efficient, specialized solution procedures. In particular, we focus on a special case of the partial-sum dispersion model defined without the propulsion factors. It should be noted that this version of the partial-sum dispersion problem is still very general because the four classic dispersion models are defined without the propulsion factors and therefore are its special cases. To avoid ambiguity, we refer to the general partial-sum dispersion problem as the *weighted* partial-sum dispersion problem and the version without propulsion factors as the *unweighted* partial-sum dispersion problem. We demonstrate that a compact and efficient formulation of the unweighted partial-sum dispersion model can be developed by using a linear program by Ogryczak and Tamir (2003) twice within the model formulation. Our computational experiments show that for the same problem instances solved by Lei and Church (2013), the new formulation can solve the partial dispersion problems faster in a majority of the cases. Moreover, we develop and present an interchange heuristic inspired by Teitz and Bart (1968) in conjunction with a branch and bound search, which are orders of magnitude faster than the integer linear programming approach. As will be shown in the experiment section, the branch and bound procedure can solve the same 55-node problem instance that took the ILP formulations a week and a half in about 3 minutes, which makes the partial-sum dispersion model suitable for location analysis in day-to-day operations.

2. Model formulation

The unweighted partial-sum dispersion problem can be formally defined as follows:

> *Locate a set of p facilities such that the sum of K worst-case facility-based sum of distances is maximized, where each facility-based sum of distances is the partial sum of L smallest distances from its neighboring facilities.*

To illustrate the concept of the partial sum dispersion problem, consider the following example in Fig. 1, in which three dispersive facilities are located among five candidate sites using a partial sum dispersion metric with $K = 1$, and $L = 2$. In the example, the coordinates for candidate sites (from A to E) are labeled. Distances between sites (in Table 1) are Euclidean distance rounded to the nearest integer.