Expert Systems with Applications 39 (2012) 4306–4314

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

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

Integrated use of fuzzy c-means and convex programming for capacitated multi-facility location problem

Tarık Küçükdeniz [a,*], Alp Baray [a], Kubilay Ecerkale [b], Şakir Esnaf [a]

[a] Department of Industrial Engineering, Faculty of Engineering, Istanbul University, Avcılar, Istanbul, Turkey
[b] Aeronautics and Space Technologies Institute, Yesilyurt, Istanbul, Turkey

ARTICLE INFO

Keywords:
Fuzzy c-means
Capacitated multi-facility location problem
Convex programming

ABSTRACT

In this study a fuzzy c-means clustering algorithm based method is proposed for solving a capacitated multi-facility location problem of known demand points which are served from capacitated supply centres. It involves the integrated use of fuzzy c-means and convex programming. In fuzzy c-means, data points are allowed to belong to several clusters with different degrees of membership. This feature is used here to split demands between supply centers. The cluster number is determined by an incremental method that starts with two and designated when capacity of each cluster is sufficient for its demand. Finally, each group of cluster and each model are solved as a single facility location problem. Then each single facility location problem given by fuzzy c-means is solved by convex programming which optimizes transportation cost is used to fine-tune the facility location. Proposed method is applied to several facility location problems from OR library (Osman & Christofides, 1994) and compared with centre of gravity and particle swarm optimization based algorithms. Numerical results of an asphalt producer's real-world data in Turkey are reported. Numerical results show that the proposed approach performs better than using original fuzzy c-means, integrated use of fuzzy c-means and center of gravity methods in terms of transportation costs.

1. Introduction

The mathematical science of facility location has attracted much research in discrete and continuous optimization over nearly four decades. Investigators have focused on both algorithms and formulations in diverse settings in the private sector (e.g., industrial plants, banks, retail facilities, etc.), and the public sector (e.g., ambulances, clinics, etc.) (Revelle & Eiselt, 2005). The facility location problem is a classical, combinatorial problem to determine the number and locations of a set of facilities (warehouses, plants, machines etc.) and assign customers to these in such a way that the total cost is minimized. If an arbitrary number of customers can be connected to a facility, the problem is called uncapacitated facility location problem. If each facility has a limit on the number customers it can serve, it becomes a capacitated facility location problem (Wu, Zhang, & Zhang, 2006). Many researchers developed and proposed numerous methods to solve uncapacitated facility location problems. Ghosh (2003) described and compared several neighborhood structures used by local search to solve this problem. Greistorfer and Rego (2006) implemented filter-and-fan approach to the uncapacitated facility location problem and compared their results with metaheuristic procedures.

Different solution approaches for capacitated facility location problems have a rich literature. Tragantalerngsak, Holt, and Rönnqvist (2000) proposed a Lagrangian relaxation-based branch and bound algorithm for capacitated facility location problem. Levin and Ben-Israel (2004) developed a heuristic method for solving large-scale multi-facility location problems. In their study, they reassign customers to facilities using the heuristic of Nearest Center Reclassification. Correia and Captivo (2006) developed a Lagrangian heuristic, enhanced by tabu search or local search in order to obtain good feasible solutions for capacitated plant location problem. Gong, Gen, Yamazaki, and Xu (1997) discussed an extension of location–allocation model, which has capacity constraints and proposed a hybrid evolutionary method to solve it is based on both genetic algorithms and evolutionary strategy. They show that the proposed method is effective in finding global or near global solutions by numerical simulations.

Holmberg, Rönnqvist, and Yuan (1999) developed a branch-and-bound method, based on the Lagrangian heuristic. Pirkul and Jayaraman (1998) presented a mixed integer-programming model to achieve the least total supply chain cost by designing a system such that all customer demand is satisfied subject to limitations imposed by production capacities for plants and storage capacities of warehouses. Hsieh and Tien (2004) solved uncapacitated

* Corresponding author. Tel.: +90 555 235 1133; fax: +90 212 473 7180.
E-mail address: tkdeniz@istanbul.edu.tr (T. Küçükdeniz).

0957-4174/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2011.09.102

location–allocation problem with rectilinear distances by using a method based on Kohonen self-organizing feature maps. They treated location–allocation problems as clustering problems and they suggested that self-organizing feature maps may provide an excellent approach when generating initial solutions for other heuristic or exact algorithms. Xu and Xu (2005) presented a linear programming rounding based approach to show that linear programming rounding techniques are equally capable of solving facility location problem.

Dohn, Christensen, and Rousoe (2007) considered different geometric and random cost problem instances. Their experiment showed that 60% of the problems can be solved to optimality by solving the corresponding LP-relaxation. Klose (2000) developed a two-stage capacitated facility location problem is to find the optimal locations of depots. Feasible solutions are constructed from Lagrangian sub problems by applying simple reassignment procedures. Computational results and comparisons with some other bounds and heuristics for this problem are presented in their study. Rönnqvist, Tragantalerngsak, and Holt (1999) presented an approach based on a repeated matching algorithm, which essentially solves a series of matching problems until certain convergence criteria are satisfied. The method generates feasible solutions in each iteration in contrast to Lagrangian heuristics where problem dependent heuristics must be used to construct a feasible solution. Numerical results show that the approach produces solutions, which are of similar and often better than those produced using the best Lagrangian heuristics. Klose and Görtz (2007) employed a column generation method in order to solve a capacitated facility location problem exactly. The column generation procedure is employed within a branch-and-price algorithm for computing optimal solutions to the problem. Canel, Khumawala, Law, and Loh (2001), Baldacci, Hadjiconstantinou, and Maniezzo (2002), Loreno and Senne (2004), and Chen and Ting (2008) are also presented different approaches for capacitated facility location problems.

The broadest categories of facility location problems are continuous problems, which have infinite set of potential locations, and discrete problems have only a finite set (Seppala, 1997). We focus on the continuous capacitated multi-facility location models, in which the facility can be located at every point in the plane and its capacity is taken into consideration. A multi-facility location problem (mflp) is defined as a special clustering problem if the sets of customers served by the same facility are considered as clusters (Levin & Ben-Israel, 2004). The crisp (classical hard) clustering methods are proposed to solve capacitated discrete and continuous mflps. Franca, Sosa, and Pureza (1999) used tabu search algorithm with two neighborhood generation mechanisms. Scheuerer and Wendolsky (2006) proposed a scatter search-based heuristic approach to the capacitated clustering problem. Negresios and Palhano (2006) proposed a two-phase polynomial heuristic algorithm. They applied a log-polynomial geometric tree search and a variable neighborhood search to the problem. Baldacci et al. (2002) presented an exact algorithm for solving the capacitated *p*-median problem based on a set partitioning formulation of the problem. Ahmadi and Osman (2005) proposed a merger of Greedy Random Adaptive Search Procedure (GRASP) and Adaptive Memory Programming (AMP) into a new GRAMPS framework for the capacitated clustering problem. Their results show that GRAMPS has an efficient learning mechanism and is competitive with the existing methods in the literature.

In contrast with crisp clustering tools, fuzzy clustering allows gradual memberships of data points to clusters in [0,1] (Döring, Lesot, & Kruse, 2006; Hu & Sheu, 2003). The fuzzy c-means clustering algorithm (Bezdek, 1981) is the most popular and its extensions have many successful applications involving finance (Turksen & Esnaf 1997a), medicine (Esnaf, 1998), MR images (Cheng, Goldgof, & Hall, 1998; Eschrich, Wei Ke, Hall, & Goldgof, 2003), inventory planning (Turksen & Esnaf, 1997b), cellular manufacturing (Ozcakar & Esnaf, 1998), acoustics (Bharitkar & Kyriakis, 2001), upwelling prediction in ocean (Nascimento, Sousa, Casimiro, & Boutov, 2005) and face recognition (Lu, Yuan, & Yahagi, 2006).

However, there are lots of applications of fuzzy c-means in different areas as mentioned above, the number of studies which employ fuzzy c-means for solving continuous uncapacitated mflps is very limited; see, e.g., Chepoi and Dimitrescu (1999), Zalik (2006), Ayoub, Martins, Wang, Seki, and Naka (2007), Esnaf and Kucukdeniz (2009). There is also single study which uses fuzzy c-means for solving continuous capacitated mflps by Esnaf, Kalkancı, and Kucukdeniz (2007).

In this study fuzzy c-means clustering algorithm- based method is proposed for solving a continuous capacitated multi-facility location problem of known demand points which are served from capacitated supply centres. As mentioned above in fuzzy c-means data points are allowed to belong to several clusters with different degrees of membership. This feature is used here to split demands between supply centers. The cluster number is determined by an incremental method, which starts with two and designated when capacity of each cluster is sufficient for its demand. Finally, each group of cluster and each model are solved as a single facility location problem. Convex programming model, which optimizes transportation cost, is used to fine tune the facility location. Numerical results of an asphalt producer's real-world data in Turkey are reported. Compared to previous literature the proposed model is the first one that combines the fuzzy c-means algorithm and the convex programming model to solve capacitated mflp.

The remainder of this paper organized as follows. Procedures of methodology and the proposed combined method are presented in Section 2. Experimental results are explained in Section 3. A real case study with numerical results generated using the proposed method and comparisons are summarized in Section 4. Finally, the concluding remarks are presented in Section 5.

2. Methodology

This methodology aims to determine optimal plant location and customer–plant assignments in respect to minimization of transportation costs. Initially, customers are grouped by fuzzy c-means clustering algorithm. Despite crisp assignments of customers to clusters which is the case when conventional clustering techniques used, fuzzy clustering gives fuzzy membership values of customers for each plant that is represented by a specific cluster. These membership values are used to determine the proportion of demand of the customer in each cluster for preassignment.

Final assignments are made by a three-step procedure. At the first step a predetermined threshold value updates the membership values in order to prevent small shipments over long distances. The second step is to allocate demand of customers to plants according to the updated membership values until the capacity of the plants are exhausted. The third step is for handling the overload capacity, if necessary.

Finally each cluster is solved as a single facility location problem. The convex programming, which optimizes transportation costs, is employed to fine-tune the facility location. The flowchart for the general mechanism of the proposed methodology is illustrated in Figs. 1 and 2.

A main performance criterion for the model is total transportation cost. In this study transportation costs include shipments of goods from plant to customers. By minimizing total transportation cost, this model also gives the customer–plant assignments. These assignments play an important role for logistical decisions made by