Computers & Operations Research 55 (2015) 23–35

Contents lists available at ScienceDirect

Computers & Operations Research

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

A quadtree-based allocation method for a class of large discrete Euclidean location problems ☆

Said Salhi [a], Chandra Ade Irawan [b,c,*]

[a] Centre for Logistics & Heuristic Optimization (CLHO), Kent Business School, University of Kent, Canterbury, Kent CT2 7PE, UK
[b] Centre for Operational Research and Logistics (CORL), Department of Mathematics, University of Portsmouth, Lion Gate Building, Lion Terrace, Portsmouth PO1 3HF, UK
[c] Department of Industrial Engineering, Institut Teknologi Nasional, Bandung, 40124, Indonesia

ARTICLE INFO

Keywords:
Allocation method
Quadtree method
p-median and *p*-center problems
Aggregation

ABSTRACT

A special data compression approach using a quadtree-based method is proposed for allocating very large demand points to their nearest facilities while eliminating aggregation error. This allocation procedure is shown to be extremely effective when solving very large facility location problems in the Euclidian space. Our method basically aggregates demand points where it eliminates aggregation-based allocation error, and disaggregates them if necessary. The method is assessed first on the allocation problems and then embedded into the search for solving a class of discrete facility location problems namely the *p*-median and the vertex *p*-center problems. We use randomly generated and TSP datasets for testing our method. The results of the experiments show that the quadtree-based approach is very effective in reducing the computing time for this class of location problems.

1. Introduction

It is a common practice when dealing with large location problems to aggregate demand points known as Basic Spatial Units (BSUs) into a small number of Aggregated Spatial Units (ASUs). Such an aggregation usually leads to error due to both distance measurement and allocation. Many of the aggregation schemes are iterative processes where the allocation must be performed several times to find the best solutions having the least errors. A two-phase approach is commonly used where in the first phase an aggregated (smaller) problem is constructed by solving a clustering problem, and in the second phase the aggregated location-allocation problem is then solved. It was noted that the design of aggregation schemes that minimize aggregation error is itself a hard problem which has not yet been solved successfully for a large number of demand points, see Francis and Lowe [11] and Francis et al. [8,9].

Hillsman and Rhoda [15] introduced three sources of error arising from demand point aggregation, known as source ABC error. Source A error happens when the distance between an ASU and a facility is applied in the model instead of the true distance between a BSU and a facility. Source B error appears in the special case when a facility is located at an ASU whereas source C error occurs when a BSU is assigned to the wrong facility. Several schemes are introduced to reduce or eliminate these types of aggregation error which are usually grouped under two categories. These include data manipulation and aggregation design which are briefly described next. For more details on aggregation methods for location problems, see Francis et al. [9].

1.1. Data manipulation

Current and Schilling [5] proposed pre-processing the demand data. Their method eliminates source A and B errors by assigning the correct total weight BSU–facility distance to each ASU–facility cell in the weighted distance matrix. The method does not address source C error. To eliminate source C error, Hodgson and Neuman [18] utilise continuous space, a set of Voronoi polygons and a GIS overlay procedure to "aggregate on the fly". Their method though is successful in addressing source C error, it fails to eliminate source A and B errors. Hodgson et al. [19] introduced a new type of error known as source D error. This occurs if some of the BSU locations happen to be at the potential sites. Bowerman et al. [4] introduced a demand portioning method that applies the Current and Schilling [5] approach to eliminate source A and B errors while producing ASUs on the fly when using a vertex interchange procedure to eliminate source C error. Hodgson and Hewko [17] studied aggregation and surrogation errors for the *p*-median problem using

☆This study was mostly conducted while the second author was at the University of Kent.

* Corresponding author.
E-mail address: chandra.irawan@port.ac.uk (C.A. Irawan).

http://dx.doi.org/10.1016/j.cor.2014.10.002

Edmonton, Canada data. The authors showed that the surrogation error was more a serious problem than the aggregation error.

1.2. Aggregation zone design

Francis and Lowe [11], Francis et al. [7,10,8,9], and Andersson et al. [2] dealt with aggregation error by developing aggregation zones for which error bounds can be determined. Their methods established rectangular zones which can be long and narrow, and hence prone to aggregation errors. Erkut and Bozkaya [6] empirically evaluated some aggregation methods. A primal-dual VNS metaheuristic for large p-median clustering problems was proposed by Hansen et al. [13] where a Reduced VNS is used to get good initial solutions which are then fed into a VNS with decomposition. Qi and Shen [23] investigated the worst-case analysis of demand point aggregation for the Euclidean p-median problem on the plane. García et al. [12] developed an alternative covering based formulation which has a small subset of constraints and variables. This method is shown to be more efficient especially when p is relatively large. Avella et al. [3] proposed an aggregation heuristic based on Lagrangean relaxation for large scale p-median problem that produced excellent results. Very recently Irawan and Salhi [20] and Irawan et al. [21] developed a multi-phase approach by solving a series of subproblems either optimally or heuristically where the obtained facility locations are then used as promising potential sites. Competitive results were generated when compared to the best known solutions. For more details on aggregation error measurements and papers dealing with aggregation to location problems, the reader will find the excellent survey paper by Francis et al. [9] informative and very valuable. The authors also point out effective as well as ineffective errors measures.

The process of determining an aggregation scheme with a minimum error is an NP-hard problem, see Francis and Lowe [11]. This difficulty has led us to develop a method where we do not aggregate demands by designing a general aggregation of demand points but we conduct aggregation with reference to a specific set of given facilities. In the location-allocation context, the method would aggregate demand relative to p trial facilities as they arise during the search as usually applied in heuristics and metaheuristics.

The main contribution of this paper is the development of an effective quadtree method (QM) used for allocating the demand points to their nearest facility when solving a class of large Euclidean discrete location problems. This allocation technique could easily be incorporated in those recent powerful algorithms for large-scale location problems as this mechanism could enhance their efficiency even further.

The paper is organized as follows. A brief on the quadtree method is given in Section 2 followed by a quadtree-based methodology in Section 3. The computational results, comparing QM against the classical allocation methods, are presented in the fourth section. The integration of QM in solving both the discrete p-median and the p-center problems is attempted in the fifth section. The last section provides a summary of our findings and highlights some suggestions for future research.

2. A brief on the quadtree method

This section demonstrates an efficient aggregation scheme that eliminates all types of allocation aggregation errors. The main idea of the scheme is adopted from the presentation given at the INFORMS conference in Montreal by Hodgson and Salhi [16]. The method utilized a spatial data compression, known as quadtrees [24], to partition the study area. The demand points could obviously also be partitioned by Voronoi polygons. Fig. 1 shows a Voronoi polygons scheme with a number of demand points and three facilities ($p=3$).

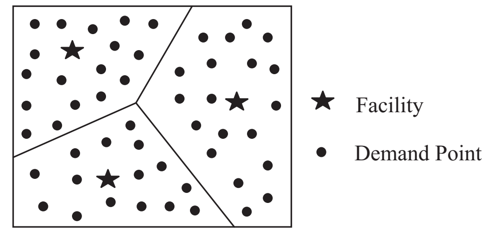

Fig. 1. A Voronoi polygons scheme.

Hodgson and Neuman [18] used this method to eliminate source C error. Noaves et al. [22] also adopted Voronoi polygons for solving continuous location-districting problems.

Though the Voronoi polygons-based allocation can be an efficient method to allocate demand points to facilities, one of the limitations is that for each new set of facilities, the new set of polygons must be generated which can be time consuming. The quadtree data structure, inspired from a raster Geographic Information System (GIS), is adapted to overcome this difficulty. The heart of QM is to pre-generate an appropriate set of common polygons with which we can systematically allocate spatial grouping of demand points to their common closest facility until all demand points have been allocated. A hierarchical organization of successively generating smaller spatial groupings is required to eliminate all aggregation errors.

A map is partitioned by raster GIS into a tessellation of square grid cells called pixels. Each pixel has its attributes, usually by assigning a number. For example a land use map might utilise 1 for green area, 2 for water, 0 for no data, and so on. Fig. 2 shows an illustration of a quadtree system where a raster grid is partitioned into a hierarchy (tree) of quadrants.

The quadtree system first partitions the map into four quadrants, each quadrant assigned with a single digit between 0 and 3. Then, each of these quadrants is then partitioned into four quadrants, each address with a second such digit. This procedure continues until a certain number of levels where each successive partition is assigned its corresponding digit between 0 and 3. Fig. 2 also presents the numbering system and the quadtree partitioning. The figure shows that the lightly shaded patch of four grid cells is assigned 200 whereas the darker grid cell is addressed 3100.

In GIS, the quadtrees are usually utilized to capture areas with the same data characteristic or attribute. Many adjacent pixels may have the same attribute in a rasterized map. In the location-allocation problem, we develop a method that adapts the quadtree structure to capture areas with the same spatial attribute (i.e., areas that are entirely closer to one facility than to any other). The number of allocations is significantly reduced by quadtrees. Fig. 3 shows how the quadtree structure deals with a location-allocation problem. There are 32×32 raster and each pixel is to be allocated to the closest of the three facilities. There is also the Voronoi polygon to recognize the correct allocation. Let L represent the quadtree level, with 0 denoting the original undivided study area. Table 1 shows the result of an example in Fig. 3. At level 1, the entire quadrant 0 (16×16 pixels) is closer to one facility than to any other. Its entirety can be allocated to that facility, it means that 256 pixels are aggregated and allocated accurately at once. Four level 2 quadrants (aggregations) are each allocated to the closest facility; in other word 256 pixels are accurately allocated. At level 3, sixteen quadrants assign another 256 pixels. Three quarters of the study area's 1024 pixels has now been accurately allocated by using 21 quadrants at the top three levels. Thirty two quadrants (128 pixels) are each allocated to the closest facility at level 4. Finally, at level five, 72 pixels can be allocated leaving 56 split only.

At some level, the limited amount of aggregation error remaining may be accepted by assigning all contained demand points to a