Expert Systems With Applications 44 (2016) 102–113

Contents lists available at ScienceDirect

Expert Systems With Applications

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

GRASP with path relinking for commercial districting

Roger Z. Ríos-Mercado[a,*], Hugo Jair Escalante[b]

[a] Graduate Program in Systems Engineering, Universidad Autónoma de Nuevo León (UANL), San Nicolás de los Garza, NL, Mexico
[b] Computer Science Department, Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE), Tonantzintla, Puebla, Mexico

ARTICLE INFO

Keywords:
Service industry
Districting
Metaheuristics
GRASP
Path relinking

ABSTRACT

The problem of grouping basic units into larger geographic territories subject to dispersion, connectivity, and balance requirements is addressed. The problem is motivated by a real-world application from the bottled beverage distribution industry. Typically, a dispersion function is minimized as compact territories are sought. Existing literature reveals that practically all the works on commercial districting use center-based dispersion functions. These center-based functions yield mixed-integer programming models with some nice properties; however, they have the disadvantage of being very costly to be properly evaluated when used within heuristic frameworks. This is due to the center updating operations frequently needed through the heuristic search. In this work, a more robust dispersion measure based on the diameter of the formed territories is studied. This allows a more efficient heuristic search computation. For solving this particular territory design problem, a greedy randomized adaptive search procedure (GRASP) that incorporates a novel construction procedure where territories are formed simultaneously in two main stages using different criteria is proposed. This also differs from previous literature where GRASP was used to build one territory at a time. The GRASP is further enhanced with two variants of forward-backward path relinking, namely static and dynamic. Path relinking is a sophisticated and very successful search mechanism. This idea is novel in any districting or territory design application to the best of our knowledge. The proposed algorithm and its components have been extensively evaluated over a wide set of data instances. Experimental results reveal that the construction mechanism produces feasible solutions of acceptable quality, which are improved by an effective local search procedure. In addition, empirical evidence indicate that the two path relinking strategies have a significant impact on solution quality when incorporated within the GRASP framework. The ideas and components of the developed method can be further extended to other districting problems under balancing and connectivity constraints.

1. Introduction

The territory design problem (TDP) may be viewed as the problem of grouping small geographic basic units (BUs) into larger geographic clusters, called territories, in a way that the territories are acceptable (or optimal) according to relevant planning criteria. Territory design or districting has a broad range of applications such as political districting, sales territory design, school districting, power districting, and public services, to name a few. The reader can find in the works of Kalcsics, Nickel, and Schröder (2005) and Duque, Ramos, and Suriñach (2007) state of the art surveys on models, algorithms, and applications to districting problems.

The problem addressed in this paper is a commercial territory design problem (CTDP) motivated by a real-world application from the bottled beverage distribution industry. The problem, introduced by Ríos-Mercado and Fernández (2009), considers finding a design of p territories with minimum dispersion subject to planning requirements such as exclusive BU-to-territory assignment, territory connectivity, and territory balancing with respect to three BU attributes: number of customers, product demand, and workload.

An important criterion in territory design problems is compactness. Typically this is achieved by minimizing a dispersion function. In commercial territory design, several models based on dispersion functions from the well-known p-center and p-median location problems have been studied in the past. These are center-based dispersion functions, that is, the dispersion is measured with respect to a centroid of a territory. However, there are other non-center-based measures of dispersion that can be used. Center-based functions rely heavily on the location of the centers; if the centers are "badly" located, the resulting design may cause a serious deterioration in objective function. In addition, in location problems, the centers represent a physical entity or facility that provides some service; however,

* Corresponding author. Tel.: +52 8183294000x1634.
E-mail addresses: roger@yalma.fime.uanl.mx, roger.rios@uanl.edu.mx (R.Z. Ríos-Mercado), hugojair@inaoep.mx (H.J. Escalante).

http://dx.doi.org/10.1016/j.eswa.2015.09.019
0957-4174/

in CTDPs the centers are artificially located as no facility is actually placed there, it is just a reference for the dispersion measure. These limitations motivate the study on other ways of measuring dispersion. For instance, a measure such as the diameter, which measures the longest distance between any two basic units in a territory, is a more robust function since it does not depend on a center location, providing more flexibility. Even from the algorithmic perspective, heuristic methods for tackling TDPs under center-based dispersion functions need to constantly update and recompute as centers keep moving along every time the territory suffers a change. This time-consuming task can be avoided if other measures such as the diameter are used.

In this work, we focus our study in a commercial territory design problem that seeks to minimize territory dispersion based on a diameter dispersion measure. To the best of our knowledge, this type of problem has not been addressed before in the territory design literature. Since the aim is to target large instances, we present a greedy randomized adaptive search procedure (GRASP) with path relinking for this NP-hard CTDP. The algorithm is denoted as GPR_CTDP. In our proposed GRASP we develop a procedure that builds exactly p territories at once simultaneously, that is, we start with p node seeds and start associating nodes to the seeds until all of them are assigned. By growing the territories simultaneously rather than one at a time one expects that the violation of the balancing constraints be considerably lower. In addition, we develop two path relinking (PR) strategies, one dynamic and one static, motivated by the work of Resende, Martí, Gallego, and Duarte (2010a), who successfully applied it to the max–min diversity problem. In our work, these PR strategies rely on finding a "path" between two different territory designs. To this end, an associated assignment subproblem for finding the best match between territory centers is solved. The solution to this problem provides a very nice way of generating the trajectory between two given designs. This idea is novel in any districting or territory design application to the best of our knowledge.

To assess its efficiency, the proposed GPR_CTDP with many of its components and strategies has been extensively evaluated over a wide set of data instances. We have found, for instance, that building territories simultaneously results in feasible solutions of acceptable quality. The two PR variants implemented in GPR_CTDP allowed us to obtain better solutions than those obtained when using straight local search; although, the static strategy resulted more helpful. The main algorithmic ideas incorporated in the developed algorithm can be extended so as to handle other districting problems with similar structure.

The paper is organized as follows. In Section 2 we describe the problem in detail and present a combinatorial optimization model. Section 3 gives an overview of relevant previous related work. Section 4 describes in detail the components of the proposed heuristic, and Section 5 presents the empirical evaluation of the method. We end the paper in Section 6, with some conclusions and final remarks.

2. Problem description

Let $G = (V, E)$ denote a graph where V is the set of city blocks or basic units (BUs), and E is the set of edges representing adjacency between blocks, that is, $(i, j) \in E$ if and only if BUs i and j are adjacent blocks. Let d_{ij} denote the Euclidean distance between BUs i and j, with $i, j \in V$. For each BU $i \in V$ there are three associated parameters. Let w_i^a be the value of activity a at node i, where $a = 1$ (number of customers), $a = 2$ (product demand), and $a = 3$ (workload). The number of territories is given by the parameter p. A p-partition of V is denoted by $X = (X_1, \ldots, X_p)$, where $X_k \subset V$ is called a territory of V. Let $w^a(X_k) = \sum_{i \in X_k} w_i^a$ denote the size of territory X_k with respect to activity $a \in A = \{1, 2, 3\}$ and $k \in K = \{1, \ldots, p\}$. The balancing planning requirements are modeled by introducing a user-specified tolerance parameter τ^a that measures the allowable relative deviation from the target average size μ^a, given by $\mu^a = w^a(V)/p$, for each activity $a \in A$. Another planning requirement is that all of the nodes assigned to each territory are connected by a path contained totally within the territory. In other words, each of the territories X_k must induce a connected subgraph of G. Finally, we seek to maximize territory compactness or, equivalently, minimize territory dispersion, where dispersion is given by the largest diameter over all territories, that is $\max_{k=1,\ldots,p} \max_{i,j \in X_k} \{d_{ij}\}$.

Let Π be the collection of all p-partitions of V. The combinatorial optimization model is given as follows.

Model (CTDP)

$$\min_{X \in \Pi} \quad f(X) = \max_{k \in K} \max_{i,j \in X_k} \{d_{ij}\} \tag{1}$$

$$\text{subject to} \quad \frac{w^a(X_k)}{\mu^a} \in [1 - \tau^a, 1 + \tau^a] \quad k \in K, \; a \in A \tag{2}$$

$$G_k = G(V_k, E(V_k)) \text{ is connected} \quad k \in K \tag{3}$$

Objective (1) measures territory dispersion. Constraints (2) represent the territory balance with respect to each activity measure as it establishes that the size of each territory must lie within a range (measured by tolerance parameter τ^a) around its average size. Constraints (3) guarantee the connectivity of the territories, where G_k is the graph induced in G by the set of nodes X_k. Note that there is an exponential number of such constraints.

The model can be viewed as partitioning G (the contiguity graph representing the BUs) into p connected components (contiguous districts) under the additional side constraints on balancing product demand, number of customers, and workload of each territory, and minimizing a dispersion measure of the BUs in a territory. The basic contiguity graph model for the representation of a territory divided into elementary units has been adopted in political districting (Ricca & Simeone, 2008).

3. Related work

Territory design or districting has a broad range of applications such as political districting (Bozkaya, Erkut, & Laporte, 2003; Browdy, 1990; Forman & Yue, 2003; Mehrotra, Johnson, & Nemhauser, 1998; Pukelsheim, Ricca, Simeone, Scozzari, & Serafini, 2012; Ricca & Simeone, 2008), sales territory design (Drexl & Haase, 1999; Zoltners & Sinha, 1983; 2005), school districting (Caro, Shirabe, Guignard, & Weintraub, 2004), power districting (de Assis, Franca, & Usberti, 2014; Bergey, Ragsdale, & Hoskote, 2003), and public services (Blais, Lapierre, & Laporte, 2003; D'Amico, Wang, Batta, & Rump, 2002; Muyldermans, Cattryse, Oudheusden, & Lotan, 2002), to name a few. The reader can find in the works of Kalcsics et al. (2005) and Duque et al. (2007) state of the art surveys on models, algorithms, and applications to districting problems. Zoltners and Sinha (2005) present a survey focusing on sales districting and Ricca, Scozzari, and Simeone (2013) present a survey on political districting.

Here we discuss the related work on commercial territory design. Ríos-Mercado and Fernández (2009) introduced the commercial TDP by incorporating a territory compactness criterion and a fixed number of territories p. They seek to maximize this compactness criterion subject to planning requirements such as exclusive BU-to-territory assignment, territory connectivity, and territory balancing with respect to three BU attributes: number of customers, product demand, and workload. In their work, the authors consider as a minimization function a dispersion function based on the objective function of the well-known p-Center Problem. After establishing the NP-completeness of the problem, the authors propose a Reactive GRASP for obtaining high-quality solutions to this problem. The core of their GRASP is a three-phase iterative procedure composed by a construction phase, an adjustment phase, and a local search phase. In