A Bi-objective Programming Model for Designing Compact and Balanced Territories in Commercial Districting

M. Angélica Salazar-Aguilar
Graduate Program in Systems Engineering
Universidad Autónoma de Nuevo León, Mexico
e.mail: *angy@yalma.fime.uanl.mx*

Roger Z. Ríos-Mercado
Graduate Program in Systems Engineering
Universidad Autónoma de Nuevo León, Mexico
e.mail: *roger@yalma.fime.uanl.mx*

José Luis González-Velarde
Center for Quality and Manufacturing
Tecnológico de Monterrey, Mexico
e.mail: *gonzalez.velarde@itesm.mx*

December 2009
Revised: May 2010

Abstract

In this paper, we address a territory design problem arising from a beverage distribution company. We propose a bi-objective programming model where dispersion and balancing with respect to the number of customers are used as performance criteria. Constraints such as connectivity and balancing with respect to sales volume are considered in the model. Most of the work in territory design has been developed for single-objective models. A very few works have addressed multi-objective territory design problems. To the best of our knowledge, this is the first multi-objective approach for this commercial territory design problem, and in particular, for territory design with connectivity constraints. In this paper, we introduce a bi-objective programming model for this problem and apply an improved ε-constraint method for generating the optimal Pareto front, based on a recent technique by Ehrgott and Ruzika to assure efficient solutions. Empirical evidence over a variety of instances shows that the improved method is well suited for finding optimal Pareto fronts at no extra computational effort than the traditional method. For this problem, the improved method finds practically the same fronts than those found by the traditional ε-constraint method. In addition, we observe that when the firm reduces the tolerance in the imbalance of sales volume the efficient fronts change and when the number of territories increases, the balance with respect to the number of customers becomes harder to achieve.

Keywords: Pareto front; improved ε-constraint method; bi-objective programming; territory design.

1 Introduction

In general, distribution firms have complex product distribution networks which are formed by thousands of sales points. In this industry there are many interesting problems from the logistic point of view that may appear in different stages of the decision process. For instance, when a firm is starting, a first problem could be where to locate the warehouses and/or distribution centers. After that, in order to provide efficient service and to reduce the total costs (i.e., production, stock, and distribution costs) some questions such as how many products need to be produced, and how to deliver the products to the final customer, need to be answered. This work is focused on the study of a problem that arises in a stage previous to the product routing and is motivated by a real-world application from a beverage distribution firm in the city of Monterrey, Mexico. The firm wants to divide the total number of customers into a specific number of groups according to some planning criteria. This partition has the objective of giving support to the decision maker when she or he designs the distribution routes and when she or he makes the workload distribution. In addition, the partition permits a more efficient management of marketing offers as it reduces the number of unsatisfied customers by applying special offers in each territory. This means that we are contributing to better route design during the routing process due to the compactness (minimum dispersion) of the territories. In addition, we provide support to the decision maker for elaborating the marketing plan and for making the best workload and resource distribution. The latter is possible because the territories are balanced with respect to both number of customers and sales volume.

This problem belongs to the family of districting problems. There has been a significant amount of work in the territory design literature addressing many different kinds of applications such as political, sales, school, services, and commercial districting, to name the most signicant. Among the most relevant works one can find Hess et al. [9], Fleischmann and Paraschis [6], Hojati [10], Garfinkel and Nemhauser [7], Mehrotra, Johnson, and Nemhauser [12], Bozkaya et al. [3], Kalcsics et al. [11]. In practically all of these works, the authors consider single-objective models. Among the very few works dealing with multi-objective districting problems we find Bowerman et al. [2], Scott et al. [16], Guo et al. [8], Bong and Wang [1], Tavares et al. [18], and Ricca and Simeone [13].

Bowerman, Hall, and Calamai [2] present a multiobjective approach for solving a school bus routing problem. They proposed a heuristic technique that at first it groups students into clustering using a multiobjective districting algorithm. After that, a school bus route and the bus stops for each cluster are generated by using a combination of a set covering procedure and a traveling salesman problem procedure. They report experimental results for a real-world instance in Wellington County, Ontario. The districting algorithm considers four objectives: minimizing the number of routes, minimizing the length of the routes, load balancing, and compactness of the routes. The three last criteria are placed in a weighted objective function where the number of routes is the dominant objective, i.e., a solution with fewer routes is always favored over a solution with more. Different plans were designed using different set of weights over the optimization criteria.

Scott, Cromley, and Cromley [16] make a multiobjective analysis of school districting in a case study from Connecticut, USA. They propose a mixed-integer goal programming model where the goal constraints are to minimize disparities in: minority enrollments, grand-list/student ratios, student-teacher ratios, and overall enrollment. The number of districts is not fixed and the contiguity criterion is not formulated in an explicit way. Experimental work using different weighting scenarios reveals that the traditional distance-minimizing or transportation-minimizing objectives are in conflict with all other aims of equity and quality of educational opportunities.

Guo, Trinidad, and Smith [8] propose a multi-objective zoning and aggregation tool (MOZART). MOZART is an integration of a graph partitioning engine with a Geographic Information System (GIS) through a graphical user interface. They illustrate the performance of MOZART by solving two zoning problems from three government local areas in Victoria: Kingston, Bayside, and Glen Eira. The first part of their experimental work is done by taking into account a single objective of equality in population size. In contrast, in the second part of their experimental work, both equity in population and compactness are treated as objective functions. They report a case with 577 census collection districts and 20 zones. The inclusion of compactness as the second zoning objective yields zones with better shapes.

Bong and Wang [1] present a multi-objective hybrid metaheuristic approach for a GIS-based spatial zoning model. Their heuristic procedure is a combination of tabu search and scatter search. They show the procedure performance by solving a political districting problem with 55 basic units and 3 districts. Equity in population, compactness, and socio-economic homogeneity are treated as objectives.

Tavares et al. [18] study a multiobjective public service districting problem. They consider multiple criteria such as location of the zone with respect to the network, mobility structure within a zone, zone corresponding to administrative structures, centers of attraction in the zone, social nature and geographical nature. They propose an evolutionary algorithm with local search and apply it to a real-world case of the Paris region public transportation. They discuss results for bi-objective cases considering different criteria combination.

Ricca and Simeone [13] address a multiple criteria political districting problem. Such criteria are connectivity, population equality, compactness and conformity to administrative boundaries. They transform the multi-objective model into a single objective model which is a convex combination of three objective functions (inequality, noncompactness and nonconformity to administrative boundaries); connectivity is considered as a constraint. They compare the behavior of four local search metaheuristics: descent, tabu search, simulated annealing, and old bachelor acceptance. The application is performed over a sample of five Italian regions where old bachelor acceptance produces the best results in most of the cases. The state of the art on territory design reveals the following facts.

Very few works address multi-objective models and all of these are basically heuristic techniques for obtaining approximate Pareto fronts. To the best of our knowledge our work is the first to provide a method for obtaining efficient fronts to bi-objective territory design problems. Single-objective versions of the commercial territory design problem addressed in this work are due to Ríos-Mercado and Fernández [14], Caballero-Hernández et al. [4], and Segura-Ramiro et al. [17]. In particular, our work can be seen as the bi-objective extension to the model developed in [17].

Our work comprises both the development of a bi-objective optimization model and an exact optimization procedure for finding efficient solutions in the sense of Pareto. The solution procedure is based on one of the most important scalarization techniques in multi-objective programming, the ε-constraint method. We implement two alternatives of this method: the traditional ε-constraint method (εCM) which guarantees obtaining weakly efficient solutions and a modified version of the ε-constraint method (IεCM) in which we include slack variables to guarantee efficient solutions. The last technique was recently proposed by Ehrgott and Ruzika [5] in the improved ε-constraint method. Our computational work reveals that the IεCM finds practically the same fronts than those found by the εCM method over all instances tested at no extra computational effort.

The rest of the paper is organized as follows. Section 2 provides a detailed description of the problem. In Section 3, the bi-objective programming model is introduced. Section 4 describes the solution method. Experimental work is discussed in Section 5 and finally we present some conclusions in Section 6.

2 Problem Description

Given a set V of city blocks (basic units, BUs), the firm wishes to partition this set into a fixed number (p) of disjoint territories that are suitable according to some planning criteria. The territories need to be balanced with respect to two different activity measures, number of customers and sales volume. Additionally, each territory has to be connected, such that the set of BUs belonging to the same territory should induce a connected subgraph. Territory compactness is required to guarantee that customers within a territory are relatively close to each other. Compactness and balance with respect to the number of customers are the most important criteria identified by the firm. In our optimization model these criteria are considered as objective functions and the remaining criteria are treated as constraints.

Let $G = (V, E)$, where V is the set of nodes (BUs) and E is the set of edges that represent adjacency between BUs (city blocks). An arc connecting nodes i and j exists if i and j are adjacent BUs. Multiple attributes like geographical coordinates (c_x, c_y), number of customers and sales volume are associated with each node $j \in V$. In particular, the firm wishes perfect balance among territories, which means each territory needs to have the same number of customers and the same sales volume. Due to the discrete nature of this problem, it is practically impossible to have perfectly balanced territories. Let $A = \{1, 2\}$ be the set of node activities, where 1 refers to the number of customers and 2 refers to sales volume. We define the size of territory V_k with respect to activity a as $w^{(a)}(V_k) = \sum_{i \in V_k}(w_i^{(a)})$, where $a \in A$ and $w_i^{(a)}$ is the value associated to activity a at node $i \in V$. Hence, the target value is given by $\mu^{(a)} = \sum_{j \in V} \frac{w_j^{(a)}}{p}$.

There are two ways to address balancing. In this work, we treat balancing with respect to the number of customers as an optimization criterion and balancing with respect to product demand as a constraint. This is motivated by the fact that this criterion is directly related with the number of stops that a vehicle makes during the product distribution and the firm pays special attention to this.

Another important constraint is that of territory connectivity. That is, it is desired that each individual territory be a connected subgraph of G. Thus, a good territory design is the one in which compactness and balancing with respect to the number of customers are optimized. In order to obtain an optimization model that includes all considerations given by the firm, we propose a bi-objective programming model in which two objective functions are minimized. The first objective f_1 is related to a dispersion measure, because minimizing dispersion is equivalent to maximizing compactness. The second objective f_2 is associated with the maximum deviation with respect to the target value ($\mu^{(1)}$) in the number of customers, minimizing the maximum deviation allows that the number of customers be closer to the average size. In this work, we use the objective of the p-median problem (p-MP) as a dispersion measure (f_1).

In a few words, the problem consists of finding a p-partition of V according to the specified planning criteria of balance with respect to the sales volume and connectivity, in such a way that both performance measures dispersion (f_1) and the maximum deviation with respect to the target number of customers in each territory (f_2) are minimized. We assume all parameters are known with certainty.

3 Bi-objective Programming Model

Indices and sets

n	number of blocks	
p	number of territories	
i, j	block indices; $i, j \in V = \{1, 2, ...n\}$	
a	activity index: $a \in A = \{1, 2\}$	
N^i	$= \{j \in V : (i, j) \in E \vee (j, i) \in E\}$	set of adjacent nodes to node i; $i \in V$

Parameters

$w_i^{(a)}$	value of activity a in node i; $i \in V$,$a \in A$
d_{ji}	Euclidean distance between j and i; $i, j \in V$
$\tau^{(2)}$	relative tolerance with respect to activity 2; $\tau^{(2)} \in [0, 1]$

Decision variables

$$x_{ji} = \begin{cases} 1 & \text{if a basic unit } j \text{ is assigned to territory with center in } i;\ i, j \in V \\ 0 & \text{otherwise} \end{cases}$$

With this definition, $x_{ii} = 1$ implies that i is a territory center.

Suppose $Q^i = \sum_{j \in V} w_j^{(1)} x_{ji} - \mu^{(1)} x_{ii}$ represents the unbalance with respect to the number of customers in territory with center in i, $i \in V$. So, the relative deviation in territory with center in $i \in V$ is given by

$$\left| \frac{Q^i}{\mu^{(1)}} \right| \tag{1}$$

This expression given as an absolute value can be decomposed into a positive ΔW_i^+ and a negative ΔW_i^- part as follows $\left| \frac{Q^i}{\mu^{(1)}} \right| = \Delta W_i^+ + \Delta W_i^-$ where $\frac{Q^i}{\mu^{(1)}} = \Delta W_i^+ - \Delta W_i^-$ and $\Delta W_i^+ \Delta W_i^- = 0, i \in V$. Based on this, we have the following bi-objective territory design problem (BOTDP) model.

$$\begin{aligned}
\text{BOTDP Min } f_1 &= \sum_{j \in V} \sum_{i \in V} d_{ji} x_{ji} && (2) \\
\text{Min } f_2 &= \max_{i \in V} \{\Delta W_i^+ + \Delta W_i^-\} && (3) \\
\text{Subject to:} & && \\
\Delta W_i^+ \Delta W_i^- &= 0 \quad i \in V && (4) \\
\Delta W_i^+ - \Delta W_i^- &= \frac{\sum_{j \in V} w_j^{(1)} x_{ji} - \mu^{(1)} x_{ii}}{\mu^{(1)}} \quad i \in V && (5) \\
\sum_{i \in V} x_{ii} &= p && (6) \\
\sum_{i \in V} x_{ji} &= 1 \quad \forall j \in V && (7) \\
\sum_{j \in V} w_j^{(2)} x_{ji} &\geq (1 - \tau^{(2)}) \mu^{(2)} x_{ii} \quad i \in V && (8) \\
\sum_{j \in V} w_j^{(2)} x_{ji} &\leq (1 + \tau^{(2)}) \mu^{(2)} x_{ii} \quad i \in V && (9) \\
\sum_{j \in \cup_{v \in S} (N^v \setminus S)} x_{ji} - \sum_{j \in S} x_{ji} &\geq 1 - \mid S \mid \quad i \in V; && \\
& \qquad S \subset [V \setminus (N^i \cup \{i\})] && (10) \\
x_{ji} &= \{0, 1\} \quad i, j \in V && (11) \\
\Delta W_i^+, \Delta W_i^- &\geq 0 \quad i \in V && (12)
\end{aligned}$$

Objective (2) represents the dispersion measure. In this sense, minimizing dispersion is equivalent to maximizing compactness. The second objective (3) represents the maximum deviation with respect to the target value of number of customers. Thus, balanced territories should have small deviation from the average number of customers. Constraints (4) and (5) establish the relationship of W_i^+ and W_i^- with the absolute value of $\frac{Q_i}{\mu^{(1)}}$. Constraint (6) guarantees the creation of exactly p territories. Constraints (7) guarantee that each node j is assigned to only one territory. Constraints (8)-(9) represent the territory balance with respect to sales volume as it establishes that the size of each territory must lie within a range (measured by tolerance parameter $\tau^{(2)}$) around its average size. Constraints (10) guarantee the connectivity of the territories. Note that there is an exponential number of such constraints.

Note that objective (3) is a piece-wise linear function. Therefore, BOTDP can be linearized by replacing (3) by (13) and introducing constraints given by (14). In addition, it can be shown (see Lemma 3.1) that the nonlinear constraints (4) are redundant.

$$\text{Min } f_2 = \gamma \tag{13}$$

$$\gamma \geq \Delta W_i^+ + \Delta W_i^-, \forall i \in V \tag{14}$$

The resulting bi-objective MILP is called LBOTDP. Model LBOTDP does not include the set of nonlinear constraints (4). It is because, when a feasible solution to LBOTDP is obtained, those indices l in which both ΔW_l^+ and ΔW_l^- take value different from zero can be easily identified. When this happens, it is always possible to get a feasible solution in which at least one of these ΔW_l^+ or ΔW_l^- takes a value equal to zero (see Lemma 3.1) and the new γ value, which will be equal or better than the actual γ value, is recomputed.

Lemma 3.1. *For any feasible solution $(X, \Delta W)$ of LBOTDP such that $\Delta W_l^+ > 0$ and $\Delta W_l^- > 0$ there exists a feasible solution $(\bar{X}, \Delta \bar{W})$ for LBOTDP such that $X = \bar{X}$ and $\Delta \bar{W}_l^+ \Delta \bar{W}_l^- = 0$, $l \in V$, where $f_1(X) = f_1(\bar{X})$ and $f_2(\Delta W) \geq f_2(\Delta \bar{W})$.*

Proof. Let $(X, \Delta W)$ be a feasible solution to $LBOTDP$ with corresponding objective function values given by (f_1, f_2). This will focus especially in constraints (5) and (14). For each $l \in L$ where $L = \{l \in V : \Delta W_l^+ > 0 \text{ and } \Delta W_l^- > 0\}$, there are two cases.

- Suppose $\Delta W_l^+ \geq \Delta W_l^-$. Let $\Delta \bar{W}_l^+ = \Delta W_l^+ - \Delta W_l^-$ and $\Delta \bar{W}_l^- = 0$. Clearly, $\Delta \bar{W}_l^+ - \Delta \bar{W}_l^- = \Delta W_l^+ - \Delta W_l^-$. Then, the new values $\Delta \bar{W}_l^+$ and $\Delta \bar{W}_l^-$ satisfy the constraints (14) as well, and $\Delta \bar{W}_l^- \Delta \bar{W}_l^+ = 0$

- Similarly if $\Delta W_l^+ < \Delta W_l^-$. Let $\Delta \bar{W}_l^- = \Delta W_l^- - \Delta W_l^+$ and $\Delta \bar{W}_l^+ = 0$. Again, $(\Delta \bar{W}_l^+, \Delta \bar{W}_l^-)$ is feasible.

Since, $\Delta\bar{W}_i^+ + \Delta\bar{W}_i^- \leq \Delta W_i{}^+ + \Delta W_i^-$, $\forall i$, it follows that $\bar{X}$ is equal to X and ΔW is less than $\Delta\bar{W}$. It implies that, $f_2(\Delta W) \leq f_2(\Delta\bar{W})$ and the proof is completed. □

From a practical point of view, it has been clearly established that both f_1 and f_2 are in conflict. It has been observed empirically that when attempting to reach the best possible dispersion measure the maximum deviation with respect to the target number of customers increases and viceversa. This justifies the bi-objective model.

4 Solution Procedure

Multiple techniques have appeared in the literature for solving multi-objective problems. One of the most important techniques used in multi-objective programming is the ε-constraint method. The ε-constraint method seems best suited for nonconvex problems such as the problem addressed here. In addition, current mono-objective approach [15] to this particular problem can be efficiently exploited within an ε-constraint frame. The ε-constraint method is based on a scalarization where one of the objective functions is minimized while all the other objective functions are bounded from above by means of additional constraints [5].

4.1 The ε-Constraint Model

In our implementation of the ε-constraint method we select the objective function given by (13) as the function to be bounded by an ε value (see LBOTPD$_\varepsilon$). We made this decision, because the firm has precisely defined the range of variation (associated with the maximum deviation γ) in which a solution is attractive to them. In addition, the resulting model has a better structure because it can be seen as a p-median problem with some additional constraints (capacity and connectivity). It is well-known that p-median models have a relatively good LP relaxation and this is true for our model as well. Finally, we tried to solve a model using f_2 as an objective and f_1 as a constraint and found a very bad LP relaxation and considerable higher run times. In general, those solutions with relative deviation (γ) less than or equal to 5% are attractive to them. Hence, different values around this value can be swept in an easy way. The model

$$
\begin{aligned}
\text{LBOTDP}_\varepsilon \quad & \text{Min } f_1 \\
\text{Subject to:} \quad & \\
& (5)\text{-}(12), \quad (14) \\
& \gamma \leq \quad \varepsilon \qquad (15)
\end{aligned}
$$

corresponds to the traditional ε-constraint (εCM) formulation for the LBOTDP model. The objective function f_1 is given explicitly by (2) and (15) is an upper bound of γ.

It is well known that the ε-constraint method guarantees to find weakly efficient solutions that can be efficient. However, when we have an optimal solution to LBOTDP$_\varepsilon$ is not easy to verify if this solution is an efficient solution or not. In order to eliminate this weakness, Ehrgott and Ruzika [5] introduced a modification of the traditional formulation. They incorporate nonnegative slack variables and with this modification the new ε-constraint method guarantees obtaining efficient solutions. Let LBPTDP$_\varepsilon^+$ be the modified ε-constraint formulation in our problem, where λ is a nonnegative weight.

$$\text{LBOTDP}_\varepsilon^+ \qquad \text{Min } f_1 - \lambda s \tag{16}$$

Subject to:

$$(5)\text{-}(12), \qquad (14)$$

$$\gamma + s \leq \quad \varepsilon \tag{17}$$

$$s \geq \quad 0 \tag{18}$$

The slack variables introduced in LBOTDP$_\varepsilon^+$ provide information about efficiency of a solution [5]. The main difference between LBOTDP$_\varepsilon$ and LBOTDP$_\varepsilon^+$ is that the ε-constraint in LBOTDP$_\varepsilon^+$ is always active at optimality.

4.2 Description of the ε-Constraint Procedures

In this work, our goal is to find both weakly efficient solutions and efficient solutions. The LBOTDP$_\varepsilon$ and LBOTDP$_\varepsilon^+$ formulations allow us to obtain these fronts by using different ε values. For each fixed value of ε we solve a single-objective problem LBOTDP$_\varepsilon$ or LBOTDP$_\varepsilon^+$. Note that each of these single-objective problems (LBOTDP$_\varepsilon$ and LBOTDP$_\varepsilon^+$) is NP-hard. In addition constraints (10) can not be written explicitly as there is an exponential number of them. There are few works that solve the single-objective districting problem with connectivity constraints. For instance, Garfinkel and Nemhauser [7] solve political districtig problems by implicit enumeration techniques, they reported successful solution for instances with up to 39 BUs and 7 territories. On the other hand, Mehrotra et al. [12] propose a column generation procedure for the political districting problem and they report solutions for up to 46 basic units and 6 territories. The iterative cut generation procedure for territory design problems (ICGP-TDP) [15] is an exact solution procedure developed for the single-objective commercial territory design problem. Empirical work shows that this procedure allows to solve instances with up to 200 basic units and 11 territories. The ICGP-TDP algorithm consists of iteratively solving a relaxed MILP model (relaxing the connectivity constraints), and then finding and adding violated constraints by solving an easy separation problem. When violated cuts are identified these are added to the model and the process continues with the next iteration. The iterative procedure continues until an optimal solution is obtained or when the relaxed problem is proved infeasible. Full details can be found in [15]. We adapted ICGP-TDP for both LBOTDP$_\varepsilon$ and LBOTDP$_\varepsilon^+$ formulations and we called it ε-ICGP.

There are a few multi-objective districting applications with connectivity constraints and these have been addressed by heuristic procedures [8, 16, 13]. To the best of our knowledge there are no references in the literature on multi-objective districting that provide exact efficient solutions. In our case, we can find weakly efficient and efficient solutions through ε-ICGP using LBOTDP$_\varepsilon$ and LBOTDP$_\varepsilon^+$ formulations, respectively. For each fixed value of epsilon, ICGP-TDP is called to obtain an optimal solution to the problem if it is feasible. At the end, the ε-ICGP procedure reports a set of efficient solutions.

The iterative solution procedure is described in Algorithm 1. The parameter λ, an initial ε value (ε_0), and a step length δ to compute different ε values are the input. Note that when $\lambda = 0$ is passed as argument to ε-ICGP, the associated solution method is the traditional εCM (see model LBOTDP$_\varepsilon$). However, when $\lambda > 0$ then the associated solution method is the IεCM (see model LBOTDP$_\varepsilon^+$). Algorithm ε-ICGP was coded in C++ and compiled with the Sun C++ 8.0 compiler under Solaris 9 Operating System. The ICGP-TDP procedure, that optimally solves the single-objective model for a fixed value of epsilon, makes use of the CPLEX 11.2 callable libraries (see [15] for more details).

While it is true that LBOTDP$_\varepsilon^+$ is more attractive than LBOTDP$_\varepsilon$ as it guarantees efficient solutions, we are interested on evaluating the computational effort of each model to properly assess the trade-off between solution quality and time.

5 Experimental Work

In the experimental work, randomly generated instances based on real-world data provided by the industrial partner were used. Each instance topology was generated by using the generator developed by Ríos-Mercado and Fernández [14]. In this work, the authors used historical information from the firm and obtained the data distribution associated to the number of customers and sales volume. A tolerance $\tau^{(2)} = 0.05$ with respect to sales volume was considered. Three different instance sets defined by $(n, p) \in \{(60, 4), (80, 5), (100, 6)\}$ were used. For each of these sets, 10 different instances were generated. Additionally, another set with five instances for $(150, 6)$ was generated. The time limit for ε-ICGP was set to 4 hours, λ was set to 3, and the step size was $\delta = 0.001$ for all instances. As it was mentioned before, solutions with a maximum deviation less than or equal to 5% from the average number of customers are attractive to the firm. Therefore, this value was used as the initial value of ε to bound the objective f_2. The procedure described in Algorithm 1 was used to optimize both the traditional and the improved formulations (LBOTDP$_\varepsilon$ and LBOTDP$_\varepsilon^+$, respectively).

The time required for both LBOTDP$_\varepsilon$ and LBOTDP$_\varepsilon^+$ formulations is first addressed. All

Algorithm 1 Solution Procedure ε-ICGP$(\lambda, \varepsilon_0, \delta)$

Input:

$\lambda :=$ Weight parameter

$\varepsilon_0 :=$ Initial ε value for bounding the objective given by f_2

$\delta :=$ Step size for computing the next ε value

Output:

$D^{\text{eff}}:=$ Efficient solution set

$D^{\text{eff}} \leftarrow \emptyset$, $\varepsilon \leftarrow \varepsilon_0$

while$(\varepsilon > 0)$

1. $S \leftarrow$ ICGP-TDP(λ, ε)

2. **if**(S is optimal)
   $D^{\text{eff}} \leftarrow D^{\text{eff}} \cup S$
   $\varepsilon \leftarrow \varepsilon - \delta$

3. **else**
   return D^{eff}

4. **end if**

end while

return D^{eff}

instance sets were tested using both formulations. It was observed that there was not a significative difference between these formulations with respect to the time and in most of the cases the set of solutions found through LBOTDP$_\varepsilon$ and LBOTDP$^+_\varepsilon$ optimization was the same. In other words, the stronger structure given by LBOTDP$^+_\varepsilon$ model takes about the same amount of computational effort. Note that, the optimization process over all instances tested stopped by time limit (4 hours). It is possible to find more efficient points if the time limit increases. So, when the time is relatively large, the algorithm continues until ε reaches the smallest value such that the problem has no feasible solutions.

Figures 1-5 show instances where the fronts obtained by the traditional εCM (LBOTDP$_\varepsilon$) and the IεCM (LBOTDP$^+_\varepsilon$) are practically the same.

Figure 1 goes here.
Figure 2 goes here.
Figure 3 goes here.

Ehrgott and Ruzika [5] show in their work that the traditional ε-constraint method (εCM) (in this case LBOTDP$_\varepsilon$) does not guarantee efficient solutions while the improved ε-constraint (IεCM) always guarantees this property. In this experimental work, the Pareto fronts reported by εCM and IεCM methods did not present significative variation, due to the fact that the constrained function (f_2) is a maximum relative deviation measure that tends to be more robust, e.g. less sensitive to changes, than, for instance, a function that measures unbalance as the sum of all deviations. In conclusion, the improved method takes about the same amount of time that the traditional method. Even though it found the same fronts than the ones found by the traditional method for these particular instances, it should be preferred as solution method.

Figure 4 goes here.

Figure 5 goes here.

The second part of this experimental work was carried out to analyze two situations that frequently take place in the firm. The first situation occurs when the number of vehicles in the fleet changes. Sometimes, economical resources decrease in a dramatic way such that the firm needs to reduce the number of vehicles (and employees) used for the distribution of the product. As a consequence the firm needs to modify the current territory design. On the other hand, when the firm experiments an expansion, it could make new employee contracts and introduce more vehicles in its fleet. This in turn means that the workload distribution will be affected and a new alignment of territories will be required. These situations were analyzed using the set of instances with 80 BUs and varying the number of territories. Figure 6 shows the set of efficient solutions obtained for an instance with 80 BUs and the number of territories $p \in \{5, 6, 7\}$. Obviously, the dispersion measure (f_2) decreases when the number of territories increases. However, it was observed that when p increases, the unbalance with respect to the number of customers is higher than when p decreases. This is because a few combinations of BUs allow to hold the connectivity constraints satisfied on each territory. Thus, the distribution of workload has more unbalance for large values of p. The decision maker needs to analyze these alternatives. She or he needs to determine what kind of territory design is better for the economical interests to the company. All instances tested with 80 BUs and $p \in \{5, 6, 7\}$ have the same behavior as the one shown in Figure 6. The results were obtained using the LBOTDP$_\varepsilon^+$ model, that is, they are efficient solutions.

Figure 6 goes here.

The second part of this last experiment was carried out to analyze the change in the Pareto front, when the tolerance ($\tau^{(2)}$) changes. The $(60, 4)$ instances for ($\tau^{(2)} \in \{0.05, 0.03, 0.015, 0.01\}$) using LBOTDP$_\varepsilon^+$ model were tested. For instance, Figure 7 shows different Pareto fronts obtained by optimizing the same instance using different $\tau^{(2)}$ values with the time limit set to 4 hours. It was observed that the Pareto front is the same for $\tau^{(2)} \in \{0.05, 0.03\}$. In contrast, the front changes when $\tau^{(2)} = 0.015$, observe that some points from the front of $\tau^{(2)} = 0.05$ remain in the front for $\tau^{(2)} = 0.015$ and additional efficient solutions are found within the time limit (4 hours).

Figure 7 goes here.

The Pareto front for $\tau^{(2)} = 0.01$ (Figure 7) shows the largest change with respect to the Pareto front obtained for $\tau^{(2)} = 0.05$. Observe for instance, the last three solutions with smallest f_2 (maximum deviation) in this front are really far from the fronts given by $\tau^{(2)} \in \{0.05, 0.015\}$. This illustrates how the front deteriorates as $\tau^{(2)}$ gets smaller.

6 Conclusions

In this paper we have presented a procedure for solving a bi-objective territory design problem with connectivity and balancing constraints. The problem is motivated by a real-world problem from a beverage distribution company. This is the first time in which the bi-objective version of this problem is addressed, to the best of our knowledge. Our solution procedure is based on the well known ε-constraint method and a cut generation procedure.

In the implementation of the exact solution procedure, two variants of the ε-constraint method are developed, i) the traditional method which guarantees to find weakly efficient solutions, and ii) the first modification proposed by Ehrgott and Ruzika [5] (in the improved ε-constraint method) which guarantees to find efficient solutions. In our computational work, it was observed that there is no significative difference between the time required by both LBOTDP$_\varepsilon$ and LBOTDP$_\varepsilon^+$ models. Moreover, both ε-constraint methods converged to practically the same Pareto fronts. The last is due to the fact that the function f_2 bounded by ε is a robust measure that corresponds to the relative deviation with respect to the average number of customers. Thus, even though the slack variable s takes a value different from zero, this value is so small such that the change in the objective value is not evident.

The performance of the proposed procedure is evaluated over a set of instances. It was observed that instances with up to 150 BUs and 6 territories are solved in a reasonable time. This is a significant result because in the general territory design literature exact solutions have been reported for instances of no more than 50 BUs. Note that this result is for the single objective case. As far as multiobjective territory design with connectivity constraints is concerned, there are no exact methods to the best of our knowledge.

Acknowledgements: We are very grateful to two anonymous referees whose comments and criticism helped improve the quality of this work. This work was supported by Universidad Autónoma de Nuevo León (UANL), grant NL-2006-C09-32652, the National Council of Science and Technology of Mexico under grants SEP-CONACYT 48499-Y and SEP-CONACYT 61343, and Tecnológico de Monterrey under research grant CAT128.

References

[1] C. W. Bong and Y. C. Wang. A multiobjective hybrid metaheuristic approach for GIS-based spatial zone model. *Journal of Mathematical Modelling and Algorithms*, 3(3):245–261, 2004.

[2] R. Bowerman, B. Hall, and P. Calamai. A multi-objective optimization approach to urban school bus routing: Formulation and solution method. *Transportation Research Part A*, 29(2):107–123, 1995.

[3] B. Bozkaya, E. Erkut, and G. Laporte. A tabu search heuristic and adaptive memory procedure for political districting. *European Journal of Operational Research*, 144(1):12–26, 2003.

[4] S. I. Caballero-Hernández, R. Z. Ríos-Mercado, F. López, and E. Schaeffer. Empirical evaluation of a metaheuristic for commercial territory design with joint assignment constraints. In J. E. Fernandez, S. Noriega, A. Mital, S. E. Butt, and T. K. Fredericks, editors, *Proceedings of the 12th Annual International Conference on Industrial Engineering Theory, Applications, and Practice (IJIE)*, pages 422–427, Cancun, Mexico, November 2007. ISBN: 978-0-9654506-3-8.

[5] M. Ehrgott and S. Ruzika. Improved ε-constraint method for multiobjective programming. *Journal of Optimization Theory and Applications*, 138(3):375–396, 2008.

[6] B. Fleischmann and J. N. Paraschis. Solving a large scale districting problem: A case report. *Computers & Operations Research*, 15(6):521–533, 1988.

[7] R. S. Garfinkel and G. L. Nemhauser. Solving optimal political districting by implicit enumeration techniques. *Management Science*, 16(8):B495–B508, 1970.

[8] J. Guo, G. Trinidad, and N. Smith. MOZART: A multi-objective zoning and aggregation tool. In *Proceedings of the Philippine Computing Science Congress (PCSC)*, pages 197–201, Manila, Philippines, 2000.

[9] S. W. Hess, J. B. Weaver, H. J. Siegfeldt, J. N. Whelan, and P. A. Zitlau. Nonpartisan political redistring by computer. *Operations Research*, 13(6):998–1006, 1965.

[10] M. Hojati. Optimal political districting. *Computers & Operations Research*, 23(12):1147–1161, 1996.

[11] J. Kalcsics, S. Nickel, and M. Schröder. Towards a unified territorial design approach: applications, algorithms, and GIS integration. *Top*, 13(1):1–56, 2005.

[12] A. Mehrotra, E. L. Johnson, and G. L. Nemhauser. An optimization based heuristic for political districting. *Management Science*, 44(8):1100–1113, 1998.

[13] F. Ricca and B. Simeone. Local search algorithms for political districting. *European Journal of Operational Research*, 189(3):1409–1426, 2008.

[14] R. Z. Ríos-Mercado and E. A. Fernández. A reactive GRASP for a commercial territory design problem with multiple balancing requirements. *Computers & Operations Research*, 36(3):755–776, 2009.

[15] M. A. Salazar-Aguilar, R. Z. Ríos-Mercado, and M. Cabrera-Ríos. New models for commercial territory design. Technical Report PISIS-2009-07, Graduate Program in Systems Engineering, UANL, San Nicolás de los Garza, México, December 2009.

[16] M. N. Scott, R. G. Cromley, and E. K. Cromley. Multi-objective analysis of school district regionalization alternatives in Connecticut. *The Professional Geographer*, 48(1):1–14, 1996.

[17] J. A. Segura-Ramiro, R. Z. Ríos-Mercado, A. M. Álvarez-Socarrás, and K. de Alba Romenus. A location-allocation heuristic for a territory design problem in a beverage distribution firm. In J. E. Fernandez, S. Noriega, A. Mital, S. E. Butt, and T. K. Fredericks, editors, *Proceedings of the 12th Annual International Conference on Industrial Engineering Theory, Applications, and Practice (IJIE)*, pages 428–434, Cancun, Mexico, November 2007. ISBN: 978-0-9654506-3-8.

[18] F. Tavares-Pereira, J. R. Figueira, V. Mousseau, and R. Bernard. Multiple criteria districting problems. the public transportation network pricing system of the Paris region. *Annals of Operations Research*, 154(1):69–92, 2007.

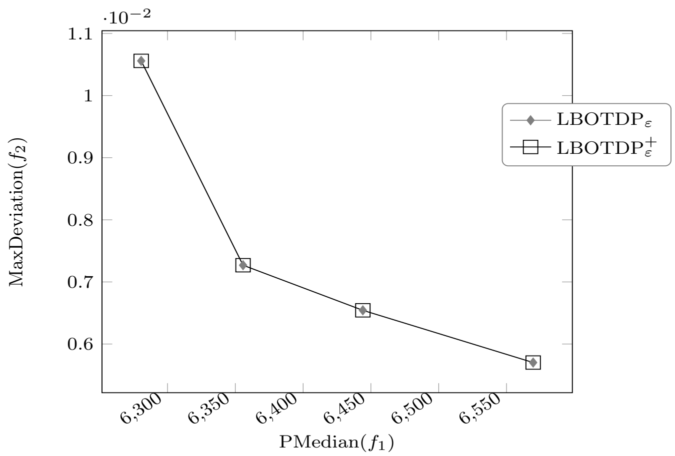

Figure 1: Comparison of LBOTDP$_\varepsilon$ and LBOTDP$_\varepsilon^+$ on an instance with 80 BUs and 5 territories.

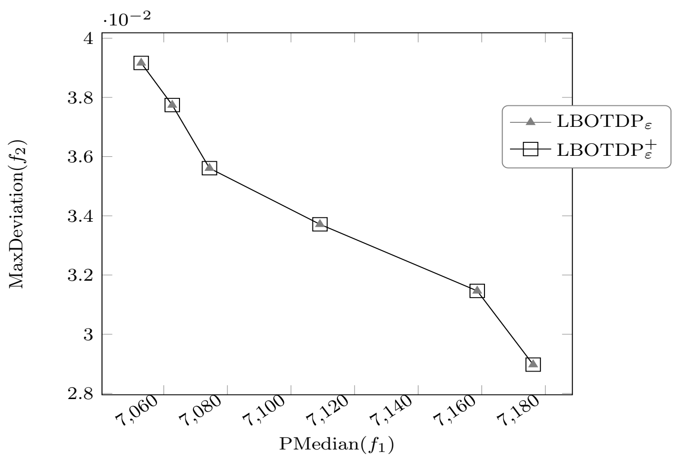

Figure 2: Comparison of LBOTDP$_\varepsilon$ and LBOTDP$_\varepsilon^+$ on an instance with 100 BUs and 6 territories.

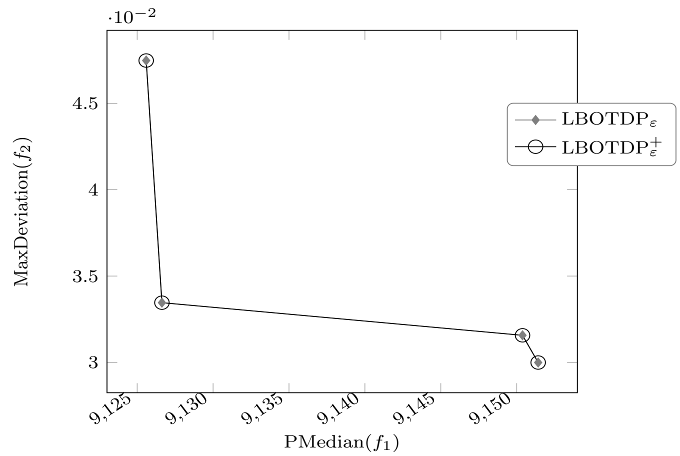

Figure 3: Comparison of LBOTDP$_\varepsilon$ and LBOTDP$_\varepsilon^+$ on an instance with 150 BUs and 6 territories.

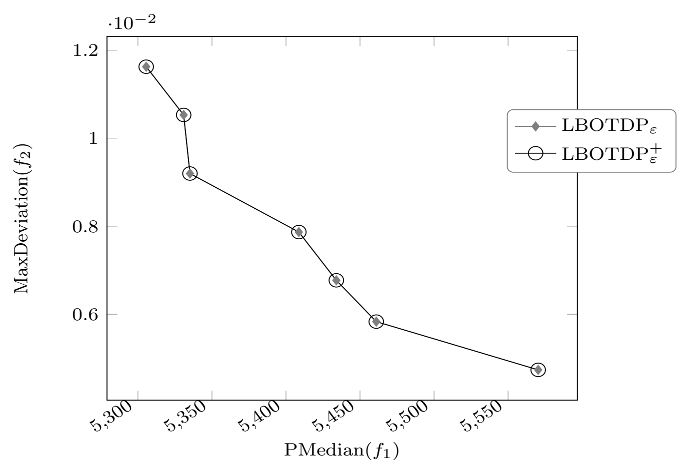

Figure 4: A) Comparison of LBOTDP$_\varepsilon$ and LBOTDP$_\varepsilon^+$ on an instance with 60 BUs and 4 territories, instance du60-01.

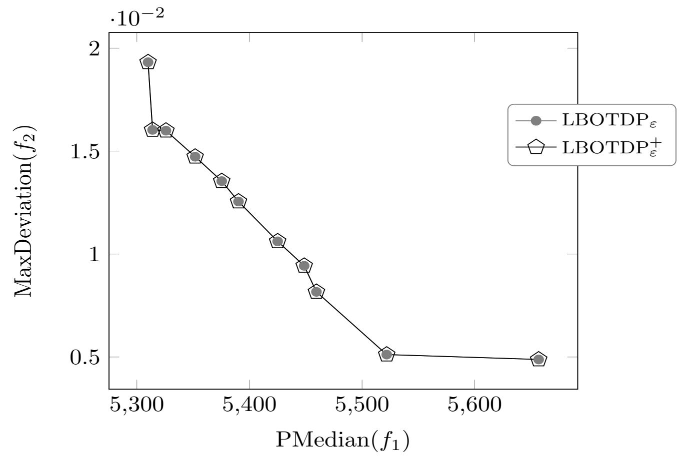

Figure 5: B) Comparison of LBOTDP$_\varepsilon$ and LBOTDP$_\varepsilon^+$ on an instance with 60 BUs and 4 territories, instance du60-08.

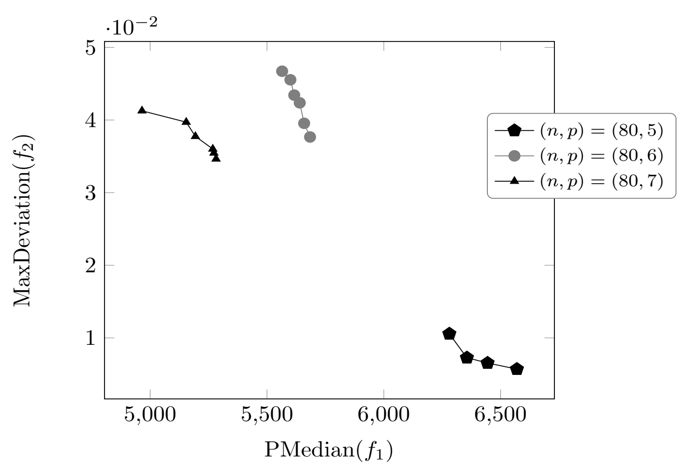

Figure 6: Changes in the efficient solutions when p changes.

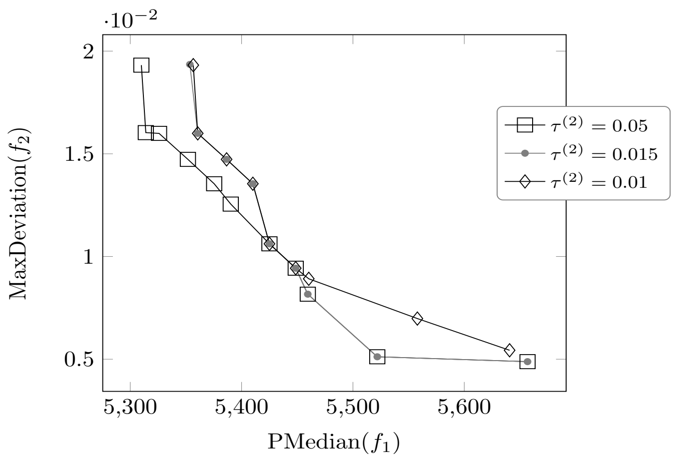

Figure 7: Comparison among Pareto fronts for different values of $\tau^{(2)}$.