Constructing Efficient Commercial Territory Design Plans for a Beverage Distribution Firm with a Bi-Objective Programming Model

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
Abstract

In this paper, we address a territory design problem arising from a beverage distribution company. This problem belongs to the family of territory design problems. 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 the sales volume are considered in the model. Most of the work in territory design have 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 a territory design with connectivity constraints. In this paper, we introduce a bi-objective mixed-integer programming model for this problem and propose an improved $\varepsilon$-constraint method for generating the optimal Pareto front. This method is based on a recent improved technique by Ehrgott and Ruzika to assure efficient solutions. Empirical evidence over a variety of instances shows that the improved method finds indeed better fronts than those found by the traditional $\varepsilon$-constraint method. This comes at no extra computational effort. In addition, we observe that when the firm reduces the tolerance in the unbalance of sales volume the efficient fronts make a dramatic change.

Keywords: Pareto front, improved $\varepsilon$-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 kind of industry there are many interesting problems from the logistic point of view and they can appear in different stages of the decision process. For instance, when a firm is starting, a first problem could be the facility location: where to install 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 in the study of a problem that arises in a stage previous to the product routing and it 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 he (she) designs the distribution routes and when he (she) makes the workload distribution. In addition, the partition permits a more efficient management of marketing offers and it reduces the number of unsatisfied customers by applying special offers in each territory. It means, we are contributing to create better route design during the routing process due to compactness (minimum dispersion) property in the territories. In addition, it provides support to the decision maker for elaborating the marketing plan and for making the best workload and resources 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 districting, sales, school districts, commercial, and services to name the most significant. Among the most relevant works one can find (Hess et al. [7], Fleischmann and Paraschis [5], Hojati [8], Garfinkel and Nemhauser [6], Mehrotra, Johnson and Nemhauser [10], Bozkaya et al. [2], Kalcsics et al. [9]). 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 (Baçao et al. [1], Tavares et al. [15], and Ricca and Simeone [11]).

Baçao et al. [1] proposed a genetic algorithm for a bi-objective political districting problem, compactness and population equality are the optimization objectives and connectivity is treated as a constraint. The algorithm was applied to a case provided by the Portuguese government. Tavares et al. [15], studied a multi-objective public service districting problem. They considered multiple criteria such as location of the zone with respect to the network, movility structure within a zone, zone corresponding to administrative structures, centers of attraction in the zone, social nature and geographical nature. They proposed an evolutionary algorithm with local search and applied it to a real-world case of the Paris region public transportation. They discussed results for bi-objective cases considering different criteria combination. Ricca and Simeone [11] addressed a multiple criteria political districting problem. Such criteria are connectivity, population equality, compactness and conformity to administrative boundaries. They transformed 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 compared the behavior of four local search metaheuristics: Descent, Tabu Search, Simulated Annealing and Old Bachelor Acceptance. The application was performed over a sample of five Italian regions and Old Bachelor Acceptance produced the best results in the majority 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 frontiers to bi-objective territory design problems. Single versions of the commercial territory design problem addressed in this work are due to (Ríos-Mercado and Fernández [12], Caballero-Hernández et al. [3], and Segura-Ramiro et al. [14]). In particular, our work can be seen as the bi-objective extension to the model developed in [14]).

Our work includes both the development of a bi-objective model and an exact optimization procedure for finding efficient solutions according to the sense of Pareto. The solution procedure is based in one of the most important scalarization techniques in multi-objective programming; the \( \varepsilon \)-constraint method. We implemented two alternatives of this method: the traditional \( \varepsilon \)-constraint method (\( \varepsilon \)CM) which guarantees obtaining weakly efficient solutions and a modified version of \( \varepsilon \)-constraint (I\( \varepsilon \)CM) in which we include slack variables to guarantee efficient solutions. The last technique was recently proposed by Ehrgott and Ruzika [4] in the improved \( \varepsilon \)-constraint method. Our computational works reveals that the I\( \varepsilon \)CM finds indeed better fronts than those found by the \( \varepsilon \)CM method. This comes at no extra computational effort. 

The rest of the paper is organized as follows. Section 2 provides a detailed description of the problem; Section 3 contains the optimization model we are proposing; 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 each of two different activity measures (number of customers and sales volume). Additionally, each territory has to be connected, so that each basic unit can be reached from other without leaving to the territory. 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 represents 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 to each node $j \in V$. In particular, the firm wishes perfect balance among territories, it means each territory needs to have the same number of customers and sales volume associated. 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. So, we define the size of territory $V_k$ with respect to activity $a$ as $w^{(a)}(V_k) = \sum_{i \in V_k} w^{(a)}_i$, where $a \in A$ and $w^{(a)}_i$ is the value associated to activity $a$ in the node $i \in V$. Hence, the target value is given by $\mu^{(a)} = \sum_{j \in V} \frac{w^{(a)}_j}{p}$.

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

Another important constraint is that of territory connectivity. That is, it is desired that each individual territory be a connected subgraph of $G$. So, the best territory design will be that in which compactness and balancing with respect to the number of customers are reached. In order to obtain an optimization model that includes all considerations given by the firm, we propose a bi-objective mixed-integer linear 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 to the maximum deviation with respect to the target value ($\mu^{(1)}$) in the number of customers, minimizing the maximum deviation allows to be closer of the average size to the number of customers. 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 way that both performance measures dispersion ($f_1$) and the maximum deviation with respect to the target number of customers on each territory ($f_2$) are minimized. We assume all parameters are known with certainly.

3 Bi-objective Mixed-Integer Programming Model

Indices and sets

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

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

In that sense \( x_{ii} = 1 \) implies \( i \) is a territory center.

Suppose \( Q_{i}^{1} = \sum_{j \in V} w_{j}^{(1)} x_{ji} - \mu^{(1)} x_{ii} \) 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

\[
\frac{|Q_{i}^{1}|}{\mu^{(1)}} (1)
\]

This expression given as an absolute value can be decomposed into a positive \( \Delta W_{i}^{+} \) and a negative \( \Delta W_{i}^{-} \) part as follows \( \frac{Q_{i}^{1}}{\mu^{(1)}} = \Delta W_{i}^{+} + \Delta W_{i}^{-} \) where \( \Delta W_{i}^{+} = \Delta W_{i}^{+} - \Delta W_{i}^{-} \) and \( \Delta W_{i}^{+} \Delta W_{i}^{-} = 0, i \in V \). Based on this, we have the following bi-objective MILP model.

**BOTDP**

\[
\text{Min } f_{1} = \sum_{j \in V} \sum_{i \in V} d_{ji} x_{ji} \\
\text{Min } f_{2} = \max_{i \in V} \{ \Delta W_{i}^{+} + \Delta W_{i}^{-} \} (2)
\]

Subject to:

\[
\Delta W_{i}^{+} \Delta W_{i}^{-} = 0 \quad i \in V (4)
\]

\[
\Delta W_{i}^{+} - \Delta W_{i}^{-} = \sum_{j \in V} w_{j}^{(1)} x_{ji} - \mu^{(1)} x_{ii} \quad i \in V (5)
\]

\[
\sum_{i \in V} x_{ii} = p (6)
\]

\[
\sum_{i \in V} x_{ji} = 1 \quad \forall j \in V (7)
\]

\[
\sum_{j \in V} w_{j}^{(2)} x_{ji} \geq (1 - \tau^{(2)}) \mu^{(2)} x_{ii} \quad i \in V (8)
\]

\[
\sum_{j \in V} w_{j}^{(2)} x_{ji} \leq (1 + \tau^{(2)}) \mu^{(2)} x_{ii} \quad i \in V (9)
\]

\[
\sum_{j \in \cup_{s \in S(N^{i}, S)}} x_{ji} - \sum_{j \in S} x_{ji} \geq 1 - |S| \quad i \in V; S \subset [V \setminus (N^{i} \cup \{i\})] (10)
\]

\[
x_{ji} \in \{0, 1\} \quad i, j \in V (11)
\]

\[
\Delta W_{i}^{+}, \Delta W_{i}^{-} \geq 0 \quad i \in V (12)
\]

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. So, balanced territories should have small deviation with respect to the average number of customers. Constraints (4) and (5) establish the relationship with the absolute value of \( \frac{Q_{i}^{1}}{\mu^{(1)}} \). 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 the 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, as usual, there is an exponential number of such constraints.

Note that objective (3) is a piece-wise linear function. So, 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 nonlinear constraints (4) are not needed.

\[
\text{Min } f_2 = \gamma 
\]

\[
\gamma \geq \Delta W_i^+ + \Delta W_i^- , \forall i \in V 
\]

The resulting bi-objective MILP is called LBOTDP. Model LBOTDP does not include the set of nonlinear constraints \( \Delta W_i^+ \Delta W_i^- = 0, i \in V \). It is because, when we get an efficient solution of LBOTDP, we can easily identify those \( L = \{l : l \in V \} \) in which both \( \Delta W_i^+ \) and \( \Delta W_i^- \) take value different to zero. When this happens, it is always possible to get an efficient solution in which at least one of these \( \Delta W_i^+ \) or \( \Delta W_i^- \) takes a value equal to zero (see Lemma 3.1) and we only recompute the new \( \gamma \) value which will be equal or better than the actual \( \gamma \) value.

**Lemma 3.1** For any efficient solution \( X \) (in the objectives space) of LBOTDP such that \( \Delta W_i^+ > 0 \) and \( \Delta W_i^- > 0 \) there exists an efficient solution \( \bar{X} \) for LBOTDP (\( \bar{X} \) that dominates \( X \)) such \( \Delta W_i^+ \Delta W_i^- = 0, l \in V \).

**Proof.** Let \( X \) be an efficient solution of LBOTDP with corresponding objective function values given by \( (y_1, y_2) \). We will focus especially in constraints (5) and (14). For each \( l \in L \) where \( L = \{l \in V : \Delta W_i^+ > 0 \text{ and } \Delta W_i^- > 0 \} \), there are two cases.

- Suppose \( \Delta W_i^+ \geq \Delta W_i^- \). Let \( \Delta \tilde{W}_i^+ = \Delta W_i^+ - \Delta W_i^- \) and \( \Delta \tilde{W}_i^- = 0 \). Clearly, \( \Delta \tilde{W}_i^+ - \Delta W_i^- = \Delta W_i^+ - \Delta W_i^- \). Then, the new values \( \Delta \tilde{W}_i^+ \) and \( \Delta \tilde{W}_i^- \) satisfy the constraints (14), too. And \( \Delta \tilde{W}_i^+ \Delta \tilde{W}_i^- = 0 \)

- Similarly if \( \Delta W_i^+ < \Delta W_i^- \). Let \( \Delta \tilde{W}_i^- = \Delta W_i^- - \Delta W_i^+ \) and \( \Delta \tilde{W}_i^+ = 0 \). Again, \( (\Delta \tilde{W}_i^+, \Delta \tilde{W}_i^-) \) is feasible.

Since, \( \Delta W_i^+ + \Delta \tilde{W}_i^- \leq \Delta W_i^+ + \Delta W_i^- \), \( \forall i \), it follows that \( \bar{X} \) can either be equal to \( X \) or dominate \( X \).

From the 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 increases and viceversa. This justifies the bi-objective model.
4 Solution Procedure

Multiple techniques have appeared in the literature for solving the multi-objective problems. One of the most important techniques used in multi-objective programming is the $\varepsilon$-constraint method. There are certainly other techniques such as the weighted sum scalarization, for instance. However, $\varepsilon$-constraint method seem best suited for nonconvex problems such as the problem addressed here. In addition, current mono-objective approach [13] to this particular problem can be efficiently exploited within an $\varepsilon$-constraint frame. The $\varepsilon$-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 [4].

4.1 The $\varepsilon$-Constraint Model

In our implementation of the $\varepsilon$-constraint method we select the objective function given by (13) as the function to be bounded by an $\varepsilon$ value (see LBOTDP$_{\varepsilon}$). We made this decision, because the firm has precisely defined the range of variation (associated to maximum deviation $\gamma$) in which a solution is attractive to them. In addition, the resulting model has a better structure so it is more tractable. In general, those solutions with relative deviation ($\gamma$) less than or equal to 5% are attractive to them. So, in the new model LBOTDP$_{\varepsilon}$ we can sweep values around this attractive value in an easy way. Then, $\varepsilon$-constraint formulation for LBOTDP is given by LBOTDP$_{\varepsilon}$ model, the objective function $f_1$ is given explicitly by (2).

\[
\text{LBOTDP}_{\varepsilon} \quad \text{Min } f_1
\]
\[
\text{Subject to:}
\]
\[
(5)-(12),(14)
\]
\[
\gamma \leq \varepsilon
\]  

(15)

It is well known that the $\varepsilon$-constraint method guarantees the obtention of 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 [4] introduced a modification in the traditional formulation. They incorporate nonnegative slack variables and with this modification the new $\varepsilon$-constraint method guarantees obtaining efficient solutions. Let LBPTDP$_{\varepsilon}^+$ be the modified $\varepsilon$-constraint formulation in our problem, where $\lambda$ is a nonnegative weight.

\[
\text{LBOTDP}_{\varepsilon}^+ \quad \text{Min } f_1 - \lambda s
\]
\[
\text{Subject to:}
\]
\[
(5)-(12),(14)
\]
\[
\gamma + s \leq \varepsilon
\]  

(17)
\[
s \geq 0
\]  

(18)

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

4.2 Solving the Single-Objective Subproblem

In this work, our goal is to find weakly efficient or efficient solutions, LBOTDP$_\varepsilon$ and LBOTDP$_\varepsilon^+$ formulations allow us to obtain these fronts by using different $\varepsilon$ values. For each fixed value of $\varepsilon$ 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 out explicitly as there are an exponential number of them. There are a few works that address single-objective districting problems with connectivity constraints. For instance, Garfinkel and Nemhauser [6] solved political districting problems by implicit enumeration techniques, they reported successful behavior for instances with up to 39 BUs and 7 territories. On the other hand, Mehrotra et al. [10] proposed a column generation procedure for a political districting problem and reported solutions for up to 46 basic units and 6 territories. ICGP-TDP procedure [13] is an exact solution procedure developed for the single-objective commercial territory 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 an 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. We adapted ICGP-TDP for both LBOTDP$_\varepsilon$ and LBOTDP$_\varepsilon^+$ formulations and called it $\varepsilon$-ICGP.

There are a few multi-objective districting applications with connectivity constraints and these have been addressed by heuristic procedures [11, 1]. To the best our knowledge there are no references in the literature on multi-objective districting that provide efficient solutions. In our case, we can find weakly efficient solutions and efficient solutions through $\varepsilon$-ICGP using LBOTDP$_\varepsilon$ and LBOTDP$_\varepsilon^+$ formulations. It means that, for each $\varepsilon$ value we call ICGP-TDP and it obtains an optimal solution to the problem when it is feasible. At the end of the $\varepsilon$-ICGP procedure, it reports a set of weakly efficient solutions (when LBOTDP$_\varepsilon$ is used) or a set of efficient solutions which belong to the Pareto optimal (when LBOTDP$_\varepsilon^+$ is used).

The iterative solution procedure is described in Algorithm 1, when LBOTDP$_\varepsilon^+$ formulation is used, user needs to specify $\lambda > 0$ previously in ICGP-TDP. In our case we select $\lambda = 3$. Note that, when $\lambda = 0$ both formulations LBOTDP$_\varepsilon$ and LBOTDP$_\varepsilon^+$ yield the same optimal solution for each $\varepsilon$ value. Algorithm $\varepsilon$-ICGP was coded in C++ and compiled with the Sun C++ 8.0 compiler. The other ICGP-TDP procedure calls ILOG CPLEX 11.2 on its iterative optimization process (see [13] for more details).

While it is true that LBOTDP$_\varepsilon^+$ is more attractive than the model LBOTDP$_\varepsilon$ as it guarantees efficient solutions, we are interested on evaluating the computational effort of each model to properly assess the trade-off.
Algorithm 1 Solution Procedure $\varepsilon$-ICGP($\varepsilon_0$, flag, TimeLimit, $\delta$)

\textbf{Input:} ($\varepsilon_0$, flag, TimeLimit, $\delta$)

$\varepsilon_0 :=$ Initial value for bounding the objective given by $f_2$
$D^{\text{eff}} \leftarrow \emptyset :=$ Set of efficient solutions
flag := equal to 0 for LBOTDP$_\varepsilon$ and equal to 1 for LBOTDP$_\varepsilon^+$
TimeLimit := time limit to find efficient solutions
Time := accumulated time through optimization process
$\delta :=$ step to compute the next $\varepsilon$ value

\textbf{Output:} $D^{\text{eff}}$ Efficient solutions set

1: $\varepsilon = \varepsilon_0$, Time = 0
2: \textbf{while} (Time < TimeLimit) \textbf{do}
3: \hspace{1em} $S \leftarrow$ ICGP-TDP(flag, $\varepsilon$)
4: \hspace{1em} Update(Time)
5: \hspace{1em} \textbf{if} (S is optimal) \textbf{then}
6: \hspace{2em} $D^{\text{eff}} = D^{\text{eff}} \cup S$
7: \hspace{2em} $\varepsilon = (1 - \delta) f_2(S)$
8: \hspace{1em} \textbf{else}
9: \hspace{2em} \textbf{return} $D^{\text{eff}}$
10: \hspace{1em} \textbf{end if}
11: \textbf{end while}
12: \textbf{return} $D^{\text{eff}}$

5 Experimental Work

In our experimental work, randomly generated instances based on real-world data provided by the industrial partner were used. Each instance topology was randomly generated as a planar graph. We considered a tolerance $\tau(2) = 0.05$ with respect to sales volume. We generated three different instance sets defined by $(n, p) \in \{(60, 4), (80, 5), (100, 6)\}$. For each of these sets, 10 different instances were generated. Additionally, we generated another set with five instances for $(150, 6)$. The time limit for $\varepsilon$-ICGP was set to four hours and we chose $\delta = 0.01$ in order to sweep the frontier in a better way. As we comment before, solutions with maximum deviation less than or equal to 5% from the average number of customers are attractive to the firm. So, we used this value as the initial value of $\varepsilon$ 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 first part of our experimental work is carried out to analyze the optimization time required for both LBOTDP$_\varepsilon$ and LBOTDP$_\varepsilon^+$ formulations. All instances sets were tested using both formulations. We could not find significative difference between these formulations with respect to the time and in most of the cases the set of weakly efficient solutions found through LBOTDP$_\varepsilon$ and the set of efficient solutions obtained by LBOTDP$_\varepsilon^+$ optimization was the same. It means that, the stronger structure given by LBOTDP$_\varepsilon^+$ model it does not cost any significant amount of computational effort. It is important to mention sometimes the LBOTDP$_\varepsilon^+$ optimization allows us to obtain one more point than the LBOTDP$_\varepsilon$ optimization and in other cases the behavior is the
Figure 1: Instance: 80 BUs and 5 territories

opposite (Figures 2 and 3). For example, in Figure 1 an instance in the set (80, 5) is shown, the weakly efficient solutions reported for the traditional $\varepsilon$CM (LBOTDP$_\varepsilon$) and the efficient solutions obtained by I$\varepsilon$CM (LBOTDP$_{+\varepsilon}$) are the same except for one point. For both formulations, the optimization process over all instances tested stopped by time limit (4 hours). It is possible to find more efficient points if we increase the time limit. So, when there is not time limit, the optimization process continues until $\varepsilon$ reaches the smallest value such that the problem has no feasible solutions.

Ehrgott and Ruzika [4] show in their work that the traditional $\varepsilon$-constraint method ($\varepsilon$CM) (in our case LBOTDP$_\varepsilon$) does not guarantee efficient solutions while the improved $\varepsilon$-constraint (I$\varepsilon$CM) always guarantees this property. This weakness of the traditional $\varepsilon$-constraint is illustrated in two instances belonging to set (60, 4), see Figures 4 and 5.

Figure 4 shows us that the set of weakly efficient solutions reported by LBOTDP$_\varepsilon$ and the set of efficient solutions obtained by LBOTDP$_{+\varepsilon}$ present a difference when $\varepsilon$ is closer to zero. In Figure 5 most of the solutions obtained by LBOTDP$_\varepsilon$ optimization are weakly efficient. The weakly efficient solutions (LBOTDP$_\varepsilon$) are really far from the efficient solutions reported by LBOTDP$_{+\varepsilon}$ optimization.

Tables 1 and 2 show the set of values attained by $f_2$ when this objective function is bounded for different $\varepsilon$ values. Each table contains results for different instances and for both LBOTDP$_\varepsilon$ and LBOTDP$_{+\varepsilon}$ models. The weakness of the traditional $\varepsilon$ is clearly illustrated in these tables. The optimization through LBOTDP$_\varepsilon$ model requires more $\varepsilon$ values than LBOTDP$_{+\varepsilon}$ to sweep the same area in the objectives space (specifically for $f_2$). For instance, Table 2 shows that during LBOTDP$_\varepsilon$ optimization fourteen $\varepsilon$ values were tested and during LBOTDP$_{+\varepsilon}$ optimization only six different $\varepsilon$ values were needed to cover the same area (in $f_2$). In addition, the set of efficient solutions for LBOTDP$_{+\varepsilon}$ is better than or equal to the weakly efficient solutions set for LBOTDP$_\varepsilon$ like Ehrgott.
Figure 2: Instance: 100 BUs and 5 territories

Figure 3: Instance: 150 BUs and 6 territories
Figure 4: Instance A: 60 BUs and 4 territories

Figure 5: Instance B: 60 BUs and 4 territories
and Ruzika [4] show in their work.

\[
\begin{array}{c|c|c|c|c}
\text{LBOTDP}_\epsilon & \text{LBOTDP}_\epsilon^+ \\
\hline
\epsilon & f_2 & \epsilon & f_2 \\
0.05 & 0.011624 & 0.05 & 0.011624 \\
0.011508 & 0.011076 & 0.011508 & 0.011076 \\
0.010965 & 0.010528 & 0.010965 & 0.010528 \\
0.010423 & 0.009197 & 0.010423 & 0.009197 \\
0.009105 & 0.007867 & 0.009105 & 0.007867 \\
0.007788 & 0.006771 & 0.007788 & 0.006771 \\
0.006703 & 0.005832 & 0.006703 & 0.005832 \\
0.005774 & 0.005597 & 0.005774 & 0.005597 \\
0.005541 & 0.005284 & 0.005541 & 0.005284 \\
0.005231 & 0.004893 & 0.005231 & 0.004893 \\
0.004844 & 0.00458 & 0.004844 & 0.00458 \\
\end{array}
\]

Table 1: Instance A: Comparison between \( \epsilon \) and \( f_2 \) values

\[
\begin{array}{c|c|c|c|c}
\text{LBOTDP}_\epsilon & \text{LBOTDP}_\epsilon^+ \\
\hline
\epsilon & f_2 & \epsilon & f_2 \\
0.05 & 0.019316 & 0.05 & 0.016031 \\
0.019123 & 0.016031 & 0.015871 & 0.012552 \\
0.015871 & 0.015045 & 0.012426 & 0.008914 \\
0.014895 & 0.014729 & 0.008825 & 0.008044 \\
0.014582 & 0.01457 & 0.007964 & 0.005118 \\
0.014424 & 0.013542 & 0.005067 & 0.004881 \\
0.013407 & 0.012552 & & & \\
0.012426 & 0.010617 & & & \\
0.010511 & 0.010142 & & & \\
0.010041 & 0.00943 & & & \\
0.009336 & 0.008165 & & & \\
0.008083 & 0.008044 & & & \\
0.007964 & 0.005118 & & & \\
0.005067 & 0.004881 & & & \\
\end{array}
\]

Table 2: Instance B: Comparison between \( \epsilon \) and \( f_2 \) values

The second part of our experimental work was carried out to analyze two situations that frequently take place in the firm. The first situation is given for increasing or reducing the number of vehicles in the fleet. 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 product. As a consequence the firm needs to modify the current territory design. On other hand, when the firm experiments an expansion, it should make new employee contracts and introduces 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. We analyze these situations using the set of instances with 80 BUs and making a variation just in 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, we can observe that when \( p \) increases, the unbalance with respect to the number of customers is higher than when \( p \) decreases. So, the distribution of workload has more unbalance for large values of \( p \). The decision maker needs to analyze these alternatives, he (she) needs to determine what kind of territory design is better for the economical interests to the company. For instance, if the decision maker decides to buy new trucks and increase the number of employees, it yields a not equitable workload distribution. In contrast, if he (she) decides to reduce the number of trucks and the number of employee contracts too, it is possible to get a better distribution of the workload.

All instances tested with 80 BUs and \( p \in \{5, 6, 7\} \) have the same behavior shown in Figure 6. The results were obtained using LBOTDP\(_{\varepsilon}^+\) model, that is, all are efficient solutions.

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

Pareto front for \( \tau^{(2)} = 0.01 \) (Figure 7) shows the largest change with respect to the other Pareto fronts. Observe for instance, the solution with smallest \( f_1 \) (dispersion measure) in this front is really far from the frontiers given by \( \tau^{(2)} \in \{0.05, 0.015\} \). So, when the firm is more restrictive in the balance of sales volume, the solutions can be worst in the objectives space.
5.10^{-2}

\text{MaxDeviation}(f_2)

PMedian(f_1)

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

6 Conclusions

In this paper we have presented a procedure for a bi-objective territory design problem with connectivity and balancing constraints. The problem is motivated by a real-world problem from a beverage distribution firm in Monterrey, Mexico. 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 $\varepsilon$-constraint method and a cut generation procedure.

We use two different versions of the $\varepsilon$-constraint method, i) the traditional method which guarantees the obtention of weakly efficient solutions and ii) the first modification proposed by Ehrgott and Ruzika [4] (the improved $\varepsilon$-constraint) which guarantees the obtention of efficient solutions. We observed there is not significative difference between the time required by both LBOTDP$_\varepsilon$ and LBOTDP$_{\varepsilon}^+$ models. So, the modification of $\varepsilon$-constraint is highly recommended in this problem, because it guarantees efficient solutions without increasing the time.

In our work, we solved instances with up to 150 BUs and 6 territories, in all of them we considered connectivity as a constraint. To the best of our knowledge, the maximum instance size that has been reported in political districting with connectivity constraint is for instances with up to 50 BUs. Note that this result is for the single objective case, we have never seen exact results for bi-objective territory design problems with connectivity constraints.

We presented an analysis for some situations that the firm faces: i) changes in the number of territories and ii) modifications in the tolerance with respect to the sales volume. These situations yield different sets of efficient solutions and provide a wide variety of alternatives to the decision maker.

Acknowledgements: This work was supported by the National Council of Science and Technol-
ogy of Mexico under grants SEP-CONACYT 48499-Y and SEP-CONACYT 61343, Universidad Autónoma de Nuevo León under its Scientific and Technological Research Support, grant UANL-PAICYT CE012-09, and Tecnológico de Monterrey under research grant CAT128. The research of the first author has been funded by a doctoral fellowship from Universidad Autónoma de Nuevo León, grant NL-2006-C09-32652.

References


