Threshold distance versus side payment to reduce the cannibalization effect in chain expansion

Blas Pelegrín*, Pascual Fernández and Juan Diego Pelegrín
Dept. Statistics and Operatinal Research
University of Murcia, Spain

María Dolores García
Dept. of Business Administration, Law, and Social Sciences
San Antonio Catholic University of Murcia, Spain

Abstract

We deal with the store location problem for an expanding retail chain in competition with other retail chains that offer the same type of product. The aim of the expanding retail chain is profit maximization, but counteracting the possible loss in profit of the existing stores in the chain caused by the appearance of the new ones. In this paper we compare two approaches, one based on a threshold distance and another based on a side payment, to reduce the effect of cannibalization of existing stores under a delivered pricing policy in a transportation network. It is proved that optimal store locations can be found at the nodes of the network and an integer linear programming model is presented to solve the store location problem for each approach. A study with data of Spanish municipalities is presented where the percentage of profit increase after the expansion and the percentage of cannibalized profit corresponding to the optimal solutions of the two models are compared for different combinations of the threshold distance, the side payment, and the number of new stores.

Keywords: Cannibalization, chain expansion, discrete optimization, location, retailing.

1 Introduction

Store location is a strategic decision for a retail chain that competes competes with other retail chains to provide the same type of product. For of stores with the same owner, the objective of locating new stores is profit maximization of the entire chain. However, this objective may be in conflict with the objective of a specific store in the chain if the stores have different owners. This happens in the franchise industry, where a franchisor is a legal entity that owns patens trademarks, operational methods, and suppliers, and it allows others to take advantage of them under its auspices. A franchisee is someone who owns and operates one store under a license agreement granted by the franchisor. The franchisees usually pay the franchisor a fixed charge at the beginning of the agreement and

*Correspondence to: Dr. Blas Pelegrín, Facultad de Matemáticas, Universidad de Murcia, 30100 Espinardo, Murcia, Spain. E-mail: pelegrin@um.es

then a royalty fee (for instance, 7% of gross sales or profit). The franchisor is interested in maximizing the system total profit, while each franchisee aims to maximize its own profit (see (see [6]). A similar situation occurs at retail level when an expanding chain locates new stores which will compete with each other as well as with any existing store (see [2, 8]).

A store in a retail chain may lose sales if a new store in the chain is open in close proximity to the existing one. This effect, known as *cannibalization*, was first considered in franchise distribution systems (see [13, 23]) and it has been well studied in the marketing literature when a new product is introduced to an existing product line (see for instance [5, 17, 19]), but it has been almost ignored in the recent location literature. Most of the proposed models study the store location problem from a multi-objective point of view. In planar location space, market share maximization and cannibalization minimization have been studied as a bi-objective problem for a single new store. The efficient frontier of locations is determined for a gravity Huff model in [9]. GIS tools have also been used to find efficient locations in [25, 26]. Cannibalization has been considered as a secondary objective for the location of a single store in [10, 24], where the cannibalized demand is minimized in the region of optimal locations with maximum demand of the entire chain. In network location space, the multi-store location problem for an expanding chain was first studied in [13] where profit of the new stores maximization and total profit of the expanding chain maximization were considered as objectives, without concern for cannibalization. A model with three objectives is proposed in [7] to design franchise outlet networks where market share of the franchisor and franchisees are constrained by a threshold value and the number of new outlets is maximized. A single-objective location model is proposed in [4] where cannibalization is not explicitly present, but the cannibalized stores may increase its demand due to market expansion as result of the entrance of new stores. The problem has also been studied with the aim of profit maximization by considering cannibalization as a cost when chains compete on delivered pricing (see [20]) and when the customer demand is estimated by Huff-like models (see [22]). Heuristics algorithms to find the Pareto front for market share maximization and cannibalization minimization have recently been proposed in [1].

The traditional method for dealing with cannibalization has been to impose territorial restrictions in franchise systems. The most common restriction is to prohibit locating new stores within a threshold distance from any existing store (see [14]). Thus, each franchisee is given the exclusive territorial right to the full demand within the threshold distance. Such a restriction has been replaced by a threshold market share constrain in [7]. A different way to reduce cannibalization is to compensate the loss of profit of existing stores as result of chain expansion by a side payment (see [20, 22]). The aim of this paper is to compare the two approach to reduce cannibalization when the competing chains use a delivered pricing policy in a transportation network. The contribution of the paper is as follows. It is proved that optimal store locations can be found in a finite set of points if location is constrained by a threshold distance. This makes possible to formulate a new integer linear programming model for profit maximization of the expanding chain under the threshold distance constrain. The increase in profit and the cannibalized profit corresponding to the optimal locations of the new stores found by the proposed threshold distance model are compared with the ones found by the model in which the cannibalized stores are compensated. A sensitivity analysis with respect to the number of new stores, the threshold distance, and the side payment is carried out by using an illustrative example with data from Spanish municipalities.

In Section 2, basic hypothesis and notation are given, the equilibrium prices are determined, and the location problem for the expanding chain is described. In Sections 3, a discretization result for the threshold distance model is proved and the problem is formulated as an integer linear programming problem. In Sections 4, the side payment model is formulated. In Section 5, the illustrative example is solved for different values of the parameters by using the two models. Finally, some conclusions are given in Section 6.

2 The location-price problem

Let $N = (V, E, l)$ be a network, with node set $V = \{v_k : k = 1, ..., n\}$, edge set $E = \{e : e = [v_k, v_j];\ v_k, v_j \in V\}$, and $l(e)$ being the length of edge e. Distance between two points a and b in the network is measured as the length of the shortest path linking the two points and it will be denoted by $d(a, b)$ (see [21]). There is a set $M = \{1, ..., m\}$ of spatially separated market areas, so that customers in market area k are assumed to be grouped at node v_k, $k = 1, ..., m$ (see [11] for demand point aggregation). Note that the network may contain some nodes on which no market is grouped, which occurs if there are some linking nodes with no customers around. Customers demand a homogeneous product and they are served from some existing stores. These stores are owned by different chains. Without loss of generality we consider two competing chains: an expanding chain A, which wants to locate new stores, and its competitors, which are named as chain B. The competitors are supposed not to react by locating other new stores, but they can change their prices after the expansion of chain A.

The two chains compete with delivered pricing, which means that each chain offers a price in each market, pays for the transportation cost, and delivers the product to the customers. Customers buy from the chain that offers the lowest price in the market they belong to. If the two chains offer the same price, customers are indifferent to chain choosing. However, the chain with the minimum marginal delivered cost (production +transportation) can offer a lower price and get the customer demand. Thus, ties in price are broken in favour of the chain with the minimum marginal delivered cost. We consider that the demand function in each market may be different from the demand function in other markets. The marginal delivered costs at each store is supposed to be independent of the amounts delivered from the store. It is assumed that the chains use linear prices and cannot sell the product at a price below their marginal delivered costs.

The location-price problem for chain A is as follows: The decision variables are the set of locations for the new stores and the set of prices to be set in the market areas after the expansion. The objective is profit maximization of the entire chain, but taken into account the cannibalization effect of the existing stores due to the expansion.

The following notation is used:

Indices

- k index of demand nodes, $k = 1, \ldots, m$.
- i index of new store location candidates (when the set of location candidates is finite)

Data

$M = \{1, 2, ..., m\}$	set of markets.
$q_k(p)$	demand function in market k in terms of price p.
S_A	set of locations for existing stores in chain A.
S_B	set of locations for existing stores in chain B.
L	set of possible locations for the new stores.
r	number of new stores to be located.
c_x	marginal production cost at location x.
t_{xk}	marginal transportation cost from location x to market k.
$C_{xk} = c_x + t_{xk}$	marginal delivered cost (or minimum delivered price) from location x to market k.

Decision variables

X	set of locations for the new stores.
p_k	price set by the expanding chain in market k, $k \in M$, after the expansion.

Miscellaneous

$d(x, k)$	distance between location x and demand point k.
$C_k(S)$ $=$	$\min\{C_{sk} : s \in S\}$ minimum delivered cost from stores in S to market k, $S \subset \{S_A \cup S_B \cup L\}$.

Notice that L, r, c_x and t_{xk} refers to chain A (chain B does not locate any new store). $C_k(S)$ may refer to either chain A or chain B depending on which chain owns the stores in S. Since each chain is supposed not to price below its marginal delivered cost, $C_k(S)$ is the minimum price to serve market k from the stores in S. Both marginal transportation cost and marginal delivered cost can be different for chains A and B.

2.1 Price competition

Once the locations of the new stores are fixed, chains A and B will compete on price. In this subsection, we summarize some of the results on equilibrium prices shown in a previous paper [12]. Such results will be used to define the location problem for the new stores.
The profit a chain gets from market k , serving the full market at price p, is $\Pi_k(p) = q_k(p)(p - C)$, where C is the marginal delivered cost of the chain. Let $q_k(p)$ be a continuous and strictly decreasing function of $p \in [0, p_k^{max}]$, where p_k^{max} is the maximum price that customers in market k are willing to pay for the product. The *monopoly price* is defined as the optimal solution to the problem: $\max\{\Pi_k(p) : C \leq p \leq p_k^{max}\}$ if this problem has a unique optimal solution. Such unique solution exists for a variety of demand functions (as linear, quadratic, exponential, hyperbolic) for which $\Pi_k(p)$ is a quasi-concave function at p. We assume that there exists a monopoly price and it is denoted by $p_k^{mon}(C)$.

In the long-term price competition the lowest price in any market k can only be offered by the chain with the minimum marginal delivered cost. Let C_k^A and C_k^B denote the minimum marginal delivered cost from the stores in chains A and B, respectively, to market k. In order to make competition effective in each market, we assume that the competing chains are able to price below the maximum price, i.e. $\max\{C_k^A, C_k^B\} < p_k^{max}$.

i) If $C_k^A < C_k^B$, then chain A gets the full market k. The maximum profit of chain A in market k is the optimal value of the following optimization problem:

$$\max\{\Pi_k(p) = q_k(p)(p - C_k^A) : C_k^A \leq p \leq C_k^B\},$$

where p is the price offered by chain A. The optimal solution to the above problem is the equilibrium price to be offered by chain A in market k and it is determined as follows:

$$p_k^*(C_k^A) = \begin{cases} p_k^{mon}(C_k^A) & if \quad p_k^{mon}(C_k^A) \leq C_k^B \\ C_k^B & if \quad p_k^{mon}(C_k^A) > C_k^B \end{cases}$$

ii) If $C_k^A = C_k^B$, the process of price competition would lead to both chains will set the same price in market k which will be equal to the minimum marginal delivered cost. Then $p_k^*(C_k^A) = C_k^A$ and no positive profit is obtained from market k.

iii) If $C_k^A > C_k^B$, any price set by the chain in market k can be lowered by its competitors, then this market will not be captured by the chain regardless of the offered price. In order to minimize the profit that its competitors get from market k, the chain will set the price $p_k^*(C_k^A) = C_k^A$.

If the competing chains set their equilibrium prices, the profit obtained by the stores in chain A can be determined as follows. Under quite general conditions, $p_k^{mon}(C_k^A)$ is a continuous increasing function for $C_k^A \in [0, p_k^{max}]$, and it is verified that $C_k^A < p_k^{mon}(C_k^A)$. Therefore, $p_k^{mon}(C_k^A) < C_k^B$ if and only if $C_k^A < \hat{C}_k$ for some value $\hat{C}_k$. The value $\hat{C}_k$ is found by solving the equation $p_k^{mon}(\hat{C}_k) = C_k^B$ once the monopoly price for the demand function $q_k(p)$ is determined. The maximum profit in market k is then defined by the function $\overline{\Pi}_k(C_k^A) = \Pi_k(p_k^*(C_k^A))$, which is given by:

$$\overline{\Pi}_k(C_k^A) = \begin{cases} \Pi_k(p_k^{mon}(C_k^A)) & if \quad C_k^A < \hat{C}_k \\ \Pi_k(C_k^B) & if \quad \hat{C}_k \leq C_k^A < C_k^B \\ 0 & if \quad C_k^B \leq C_k^A. \end{cases}$$

Therefore, the total profit obtained by the stores in chain A is:

$$\Pi_A = \sum_{k=1}^{n} \overline{\Pi}_k(C_k^A)$$

2.2 The location problem

The location-price problem for the expanding chain can be reduced to a location problem if the competing chains set the equilibrium prices. Due to the price competition process, chain A gets a positive profit from the markets in which the minimum marginal delivered cost from its stores is less than the marginal delivered cost of any competitor. The lower the marginal delivered cost, the higher the profit obtained by the expanding chain. Then, in order to maximize the profit of the entire chain, the product will be delivered from one of the stores in chain A with the minimum marginal delivered cost.

Before the expansion, the set of profitable market areas captured by chain A is,

$$M_A = \{k \in M : C_k(S_A) < C_k(S_B)\}.$$

Since it is assumed that the competing chains set the equilibrium prices and the minimum delivered cost in market k is $C_k(S_A)$, the total profit of chain A before the expansion is:

$$\Pi_A = \sum_{k \in M_A} q_k(p_k^*(C_k(S_A)))(p_k^*(C_k(S_A)) - C_k(S_A)).$$

After the expansion, the set of markets in which chain A increases its profit depends on the location of the new stores. If X is the set of locations for the new stores in chain A, the chain will increase its profit in the following set of markets,

$$M_X = \{k \in M : C_k(X) < \min\{C_k(S_A), C_k(S_B)\}\}.$$

Then chain A obtains the same profit as the one obtained before the expansion from any market $k \in M_A \backslash M_X$, but it increases its profit in any market $k \in M_X$. Since the minimum delivered cost from the expanding chain to each one of these markets is known, the maximum profit of the entire chain after the expansion is given by the following function:

$$\begin{aligned}\Pi_A(X) = &\sum_{k \in M_A \backslash M_X} q_k(p_k^*(C_k(S_A)))(p_k^*(C_k(S_A)) - C_k(S_A)) \\ &+ \sum_{k \in M_X} q_k(p_k^*(C_k(X)))(p_k^*(C_k(X)) - C_k(X)).\end{aligned}$$

Some of the profitable markets that were served by some existing stores in chain A before the expansion will be served by the new stores after the expansion. Such markets are the ones in the set $M_{Can}(X) = M_A \bigcap M_X$. The existing stores which lose profit as a result of the expansion are called *cannibalized stores*. The total profit lost by the cannibalized stores is:

$$\Pi_{Can}(X) = \sum_{k \in M_{Can}(X)} q_k(p_k^*(C_k(S_A)))(p_k^*(C_k(S_A)) - C_k(S_A)).$$

It is assumed that chain A obtains a portion of the profit obtained by each one of its stores. Let γ be the amount of money obtained by chain A per unit of profit of the stores in the chain, $0 < \gamma < 1$. Then the location problem for the expanding chain is how to select the set X with the aim of maximizing $\gamma\Pi_A(X)$, but avoiding that $\Pi_{Can}(X)$ will be high.

In the following we study the two mentioned approach to deal with the location problem.

3 The threshold distance model

Let us consider that the owner of chain A and the owners of existing stores in chain A agree the new stores cannot be located to a distance from any existing store less than a given threshold distance. This agreement is based on the fact that the cannibalization effect can be smaller by locating the new stores far from the existing stores in chain A. Then the set of location candidates is reduced to:

$$L_D = \{x \in L : d(x, j) \geq D, \forall j \in S_A\}$$

where D is the threshold distance.

The distance function $d(x, j)$ with x in any given edge $[u, v]$ is piece linear concave for any $j \in S_A$, which implies that the function $d(x, S_A) = \min\{d(x, j) : j \in S_A\}$ with $x \in [u, v]$ is also piece linear concave. Let $T_D = \{x \in E : d(x, S_A) = D\}$. Then the set of feasible locations in $[u, v]$ for the new stores is the segment $[\overline{u}, \overline{v}] = [u, v] \cap L_D$, where $\overline{u} = u$ or $\overline{u} \in T_D$ and $\overline{v} = v$ or $\overline{v} \in T_D$ (see Fig. 1).

Then the aim of chain A is profit maximization taking into account such agreement, which implies to solve the following network optimization problem:

$$(P_D) : \max\{\gamma\Pi_A(X) : |X| = r, X \subset L_D\}$$

If $L = V \cup E$, the following property is verified.

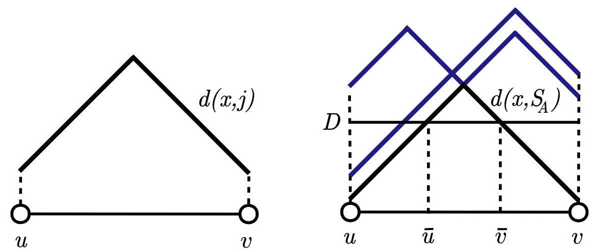


Figure 1: Distance functions and points in T_D

Property 1 *If C_{xk} is concave when x varies in any edge, then an optimal solution to problem (P_D) can be found in $V \cup T_D$.*

Proof: The objective function of problem (P_D) can be expressed as:

$$\gamma\Pi_A(X) = \gamma \sum_{k=1}^{n} \overline{\Pi}_k(C_k(S_A \cup X)).$$

Let $X = \{x_1, x_2, .., x_r\} \subset L_D$, where all locations in the X are fixed, with the exception of one location x_i for some given index i. Let x_i vary in some edge $[u, v]$. As C_{kx_i} is a concave function at x_i in $[u, v]$ and $C_k(S_A \cup X) = \min\{C_k(S_A), C_k(X \setminus x_i), C_{kx_i}\}$, then $C_k(S_A \cup X)$ is also a concave function at x_i in $[u, v]$ if the points in $X \setminus x_i$ are fixed. In [12] it is shown that $\overline{\Pi}_k(C_k)$ is a convex decreasing function at $C_k \geq 0$. From the theorem of composition of convex functions (see [3]) it is obtained that $\Pi_k(C_k(S_A \cup X))$ is convex at x_i in $[u, v]$. Since a linear combination of convex functions with positive coefficients is also convex, it follows that the profit function $\gamma\Pi_A(X)$ is convex at x_i in $[u, v]$ if the points in $X \setminus x_i$ are fixed.

Due to the convexity property of $\gamma\Pi_A(X)$ at x_i, for any $x_i \in [\overline{u}, \overline{v}]$ it is verified that:

$$\gamma\Pi_A(X)) \leq \gamma \max\{\Pi_A(\{X \setminus x_i\} \cup \overline{u}), \Pi_A(\{X \setminus x_i\} \cup \overline{v})\}.$$

Therefore, replacing each point x_i in X, where x_i is in some edge $[u, v]$, by one of the points $\overline{u}$ or $\overline{v}$ (the one with the maximum value of the objective function), we will obtain a set $\hat{X} \subset V \cup T_D$ such that $\gamma\Pi_A(X) \leq \gamma\Pi_A(\hat{X})$.Therefore, an optimal solution to problem (P_D) can be found in $V \cup T_D$. $\square$

Note that for each edge $[u, v]$ there are at most two points in $[u, v] \cap T_D$, then the set $V \cup T_D$ contains at most $2|E|$ points. Then optimal locations to problem (P_D) can be found by searching in a finite set of points on the network, if the marginal delivered cost is concave along any edge. For the motivation of the concavity assumption the reader is referred to [15, 16]. In the following we show that problem (P_D) can be solved as an integer linear programming problem when the set of location candidates is finite.

Integer linear programming formulation

For a fixed set X of store locations, the total profit obtained by the stores in chain A is given by,

$$\Pi_A(X) = \Pi_A + \Pi_{New}(X) - \Pi_{Can}(X),$$

where $\Pi_{New}(X)$ is the profit obtained by the new stores. Then maximizing the function $\gamma\Pi_A(X)$, taking into account the threshold distance constraint, is equivalent to the following maximization problem :

$$\max\{\Pi_{New}(X) - \Pi_{Can}(X) : |X| = r, X \subset L_D\}.$$

The set of markets M_A captured by chain A before the expansion can be split into the subsets:

$$M_A^1 = \{k \in M_A : p_k^{mon}(C_k(S_A)) < C_k(S_B)\}$$

$$M_A^2 = \{k \in M_A : p_k^{mon}(C_k(S_A)) \geq C_k(S_B)\}$$

The set of location candidates at which a new store can price below the price set by the existing stores in market k is:

$$L_D^k = \{i \in L_D : C_{ik} < \min\{C_k(S_A), C_k(S_B)\}\}$$

The set of markets which can be captured by the new stores after the expansion is:

$$M^* = \{k \in M : L_D^k \neq \emptyset\}$$

The set L_D^k can be split into the subsets:

$$L_D^{k1} = \{i \in L_D^k : p_k^{mon}(C_{ik}) < C_k(S_B)\}$$

$$L_D^{k2} = \{i \in L_D^k : p_k^{mon}(C_{ik}) \geq C_k(S_B)\}$$

The following variables are considered:

$$x_i = \begin{cases} 1 & \text{if a new store is located in } i \\ 0 & \text{otherwise} \end{cases} \qquad i \in L_D$$

$$y_{ik} = \begin{cases} 1 & \text{if market } k \text{ is served from } i \\ 0 & \text{otherwise} \end{cases} \qquad k \in M^*, \quad i \in L_D^k$$

$$z_k = \begin{cases} 1 & \text{if market } k \text{ is captured} \\ 0 & \text{otherwise} \end{cases} \qquad k \in M^*$$

Then problem (P_D) can be formulated as:

$$\begin{aligned}
\text{Maximize} \quad & \textstyle\sum_{k\in M^*} \sum_{i\in L_D^{k1}} q_k(p_k^{mon}(C_{ik}))(p_k^{mon}(C_{ik}) - Cik)y_{ik} \\
& + \textstyle\sum_{k\in M^*\setminus M_A^1} \sum_{i\in L_D^{k2}} q_k(C_k^B)(C_k^B - C_{ik})y_{ik} \\
& - \textstyle\sum_{k\in M^*\cap M_A^1} q_k(p_k^{mon}(C_k(S_A)))(p_k^{mon}(C_k(S_A)) - C_k(S_A)) \sum_{i\in L_D^k} y_{ik} \\
& - \textstyle\sum_{k\in M^*\cap M_A^2} q_k(C_k^B)(C_k^B - C_k(S_A)) \sum_{i\in L_D^k} y_{ik}
\end{aligned}$$

$$
\begin{array}{lllll}
\text{subject to :} & \sum_{i\in L_D^k} y_{ik} & \leq & 1, & k \in M^* & (1)\\
& y_{ik} & \leq & x_i, & k \in M^*, i \in L_D^k & (2)\\
& \sum_{i\in L_D} x_i & = & r & & (3)\\
& \sum_{i\in L_D^k} x_i & \leq & r z_k & k \in M^* & (4)\\
& \sum_{i\in L_D^k} y_{ik} & \geq & z_k & k \in M^* & (5)\\
& x_i, y_{ik}, z_k & \in & \{0,1\} & k \in M^*, i \in L_D^k & (6)
\end{array}
$$

The objective function evaluates the increment in profit of the entire chain after the expansion, which is given by the profit obtained of the new stores minus the cannibalized profit from the existing stores. Constraints (1) mean that each market $k \in M^*$ can be served from at most one of the new stores (the store with the minimum marginal delivered cost in the optimal solution). Constraints (2) imply that a variable y_{ik} may be positive only if a store is located at i. Constraint (3) represents the number of new stores to be located. Constraints (4) and (5) guarantee that k will be served by a new store if $x_i = 1$, $i \in L_k$. Finally, constraints (6) mean that all variables are constrained to be binary.

4 The side payment model

An alternative to the threshold distance agreement is that chain A compensates those stores that could be cannibalized as consequence of the expansion. The agreement here is that chain A will pay an amount δ of money to any cannibalized store per unit of lost profit. As γ is the portion of profit that the chain obtains from the stores, it is reasonable to assume that $0 < \delta \leq 1 - \gamma$.

Since chain A is interested in maximizing its profit, the best set of new store locations, taking into account the new agreement, is the optimal solution to the following optimization problem:

$$(P_{\gamma\delta}) : \max\{\gamma\Pi_A(X) - \delta\Pi_{Can}(X) : |X| = r, X \subset L\}.$$

As $\Pi_A(X) = \Pi_A + \Pi_{New}(X) - \Pi_{Can}(X)$, maximizing the function $\gamma\Pi_A(X) - \delta\Pi_{Can}(X)$ is equivalent to maximizing the function $\gamma\Pi_{New}(X) - (\gamma + \delta)\Pi_{Can}(X)$.

Integer linear programming formulation

If the set of location candidates is finite, then the previous problem can be formulated as an integer linear programming problem as follows: The set of markets M_A captured by chain A before the expansion is split into the subsets:

$$M_A^1 = \{k \in M_A : p_k^{mon}(C_k(S_A)) < C_k(S_B)\}$$
$$M_A^2 = \{k \in M_A : p_k^{mon}(C_k(S_A)) \geq C_k(S_B)\}$$

The set of location candidates at which a new store can price below the price set by the existing stores in market k is:

$$L^k = \{i \in L : C_{ik} < \min\{C_k(S_A), C_k(S_B)\}\}$$

The set of markets which can be captured by the new stores after the expansion is:

$$M^* = \{k \in M : L^k \neq \emptyset\}$$

The set L^k can be split into the subsets:

$$L^{k1} = \{i \in L^k : p_k^{mon}(C_{ik}) < C_k(S_B)\}$$
$$L^{k2} = \{i \in L^k : p_k^{mon}(C_{ik}) \geq C_k(S_B)\}$$

The following variables are considered:

$$x_i = \begin{cases} 1 & \text{if a new store is located in } i \\ 0 & \text{otherwise} \end{cases} \qquad i \in L$$

$$y_{ik} = \begin{cases} 1 & \text{if market } k \text{ is served from } i \\ 0 & \text{otherwise} \end{cases} \qquad k \in M^*, \quad i \in L^k$$

$$z_k = \begin{cases} 1 & \text{if market } k \text{ is captured} \\ 0 & \text{otherwise} \end{cases} \qquad k \in M^*$$

Then problem $(P_{\gamma\delta})$ is equivalent to:

$$\begin{aligned} \text{Maximize} \quad & \gamma \textstyle\sum_{k\in M^*} \sum_{i \in L^{k1}} q_k(p_k^{mon}(C_{ik}))(p_k^{mon}(C_{ik}) - Cik)y_{ik} \\ & +\gamma \textstyle\sum_{k \in M^* \setminus M_A^1} \sum_{i\in L^{k2}} q_k(C_k^B)(C_k^B - C_{ik})y_{ik} \\ & -(\gamma+\delta) \textstyle\sum_{k\in M^*\cap M_A^1} q_k(p_k^{mon}(C_k(S_A)))(p_k^{mon}(C_k(S_A)) - C_k(S_A)) \sum_{i\in L^k} y_{ik} \\ & -(\gamma+\delta) \textstyle\sum_{k\in M^*\cap M_A^2} q_k(C_k^B)(C_k^B - C_k(S_A)) \sum_{i \in L^k} y_{ik} \end{aligned}$$

$$\begin{array}{llllll} \text{subject to :} & \sum_{i\in L_k} y_{ik} & \leq & 1, & k \in M^* & (1) \\ & y_{ik} & \leq & x_i, & k \in M^*, i \in L^k & (2) \\ & \sum_{i \in L} x_i & = & r & & (3) \\ & \sum_{i\in L^k} x_i & \leq & r z_k & k \in M^* & (4) \\ & \sum_{i \in L^k} y_{ik} & \geq & z_k & k \in M^* & (5) \\ & x_i, y_{ik}, z_k & \in & \{0,1\} & k \in M^*, i \in L^k & (6) \end{array} .$$

Observe that the objective function is the increment in profit of the owner of chain A after the expansion. The constraints are the same as in the threshold distance model if L_D is replaced by L.

Notice that it is possible that only markets within the threshold distance are compensated (as the others may be too far away). Such markets are the ones in the set $M_A^D = \{k \in M_A : d(k, S_A) = \min\{d(k,i) : i \in S_A\} \leq D\}$. Then, $M^* \cap M_A^1$ and $M^* \cap M_A^2$ should be replaced by $M_A^D \cap M^* \cap M_A^1$ and $M_A^D \cap M^* \cap M_A^2$, respectively, in the objective function.

5 Comparison between the two models

The locations obtained by the threshold distance (TD) model are suitable if the increase in profit is high and cannibalization is small when the new stores are located at such locations. We wonder how big the threshold distance D should be taken to get suitable locations. The locations obtained by the side payment (SP) model are suitable if the increase in profit is high and δ is close to $1-\gamma$. We wonder if the increase in profit obtained with the SP model can be higher than, or equal to, the increase in profit obtained with the TD model when $\delta = 1-\gamma$. In such a case, the SP model would be preferred to the TD model.

Let X_D denote a set of optimal locations to problem (P_D). The percentage of profit increase of the owner of chain A after the expansion when the new stores are located in X_D is given by:

$$\triangle\Pi_A(X_D) = \frac{\Pi_{New}(X_D) - \Pi_{Can}(X_D)}{\Pi_A} \times 100$$

and the percentage of cannibalized profit is:

$$\triangle\Pi_{Can}(X_D) = \frac{\Pi_{Can}(X_D)}{\Pi_A} \times 100$$

Observe that neither the optimal solution nor the percentages depend on the γ value.

Let X_γ denote a set of optimal locations to problem $(P_{\gamma,\delta})$ for $\delta = 1-\gamma$. The percentage of profit increase of the owner of chain A after the expansion when the new stores are located in X_γ is given by:

$$\triangle\Pi_A(X_\gamma) = \frac{\gamma\Pi_{New}(X_\gamma) - \Pi_{Can}(X_\gamma))}{\gamma\Pi_A} \times 100$$

the percentage of cannibalized profit:

$$\triangle\Pi_{Can}(X_\gamma) = \frac{\Pi_{Can}(X_\gamma)}{\Pi_A} \times 100$$

In the following we use an illustrative example to compare those percentages for different values of the number of new stores, the threshold distance and the side payment. We will remark the values of D and γ for which $\triangle\Pi_A(X_\gamma) > \triangle\Pi_A(X_D)$.

An illustrative example

We consider the transportation network in Spain where municipalities over 10,000 inhabitants are taken as markets. These municipalities have been numbered from 1 to 615 in decreasing population size, thus $M = \{1, 2, ..., 615\}$ (see Fig. 2). The location candidates are municipalities over 20,000 inhabitants, $L = \{1, 2, ..., 314\}$ (see Fig. 3). The competing stores are located at municipalities (the nodes of the Spanish transportation network). Let us assume that there are two existing stores in chain A, being $S_A = \{42, 114\}$, and five existing stores in chain B, being $S_B = \{74, 76, 120, 122, 309\}$ (see Fig. 4). The population size and geographical coordinates of the Spanish municipalities have been taken from the web: *https://www.businessintelligence.info/varios/longitud-latitud-pueblos-espana.html*. Distances $d(i, k)$ between any pair of municipalities i and k have been approximated by using the Harversine formula, which measure the distance between two geographical points from their longitudes and latitudes [18].

We have used the following data:

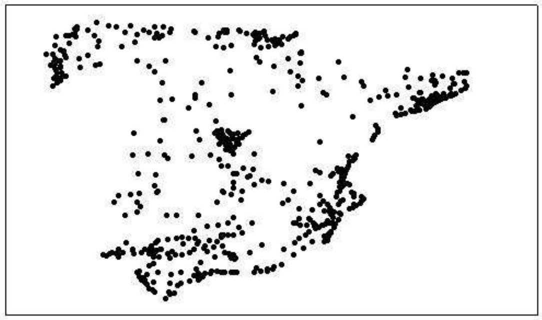


Figure 2: Demand points

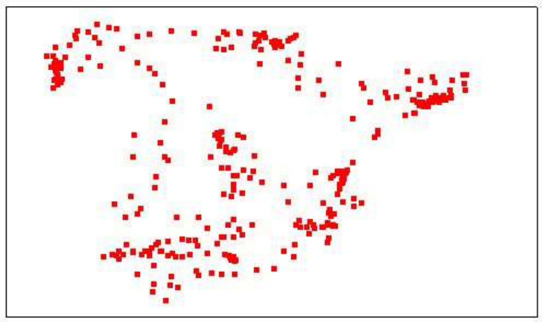


Figure 3: Location candidatess

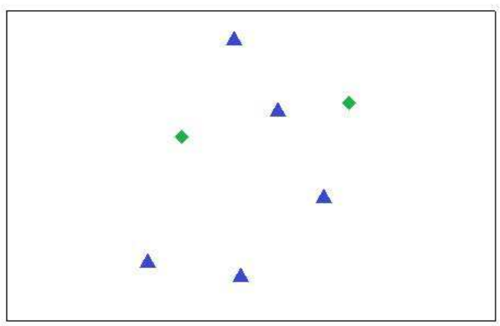


Figure 4: Stores in chain A: $\Diamond$. Stores in chain B: $\triangle$

Demand function

The maximum price to be paid for the product is 700 euros. Since the greater size of municipality, the greater demand of the product, we have taken the following linear demand function at each municipality k:

$$q_k(p) = m_k - \frac{m_k}{700} \quad p \quad , \qquad\qquad 0 \leq p \leq 700$$

where m_k is proportional to the population at municipality k,

$$m_k = \frac{1}{1000} \times \text{size of municipality } k.$$

Marginal production cost

The marginal production cost may be supposed not to be decreasing in the size of the population. The higher value of m_i for each municipality i, the higher or equal cost c_i. Thus, we have taken the following values for c_i:

m_i	$1000 \leq m_i$	$600 < m_i \leq 1000$	$300 < m_i \leq 600$
c_i	200	180	160
m_i	$100 < m_i \leq 300$	$50 \leq m_i \leq 100$	
c_i	140	120	

Marginal delivered cost

For simplicity, the marginal transportation cost t_{ik} is taken equals to the distance $d(i,k)$ between municipalities i and k. Thus, the marginal delivered cost from location i to municipality k is:

$$C_{ik} = c_i + d(i,k)$$

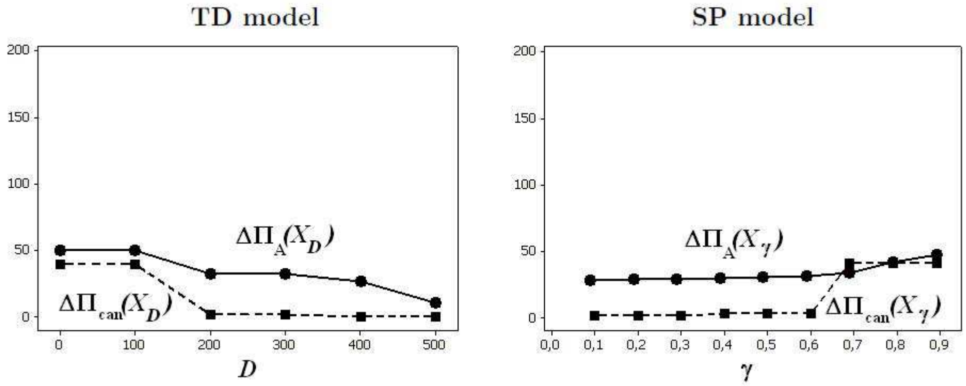
We have selected a wide range of values for the parameters D, γ and r, in order to cover a variety of possible cases. The greater value of D, the smaller candidates for locating the new stores with the threshold distance model. Since the maximum price is 700, with the previous data, if there is a store at municipality i where the marginal production cost is 200, the store could only serve markets to a distance from i less than 500 km. There are a great number of municipalities within a distance of 500 km. from any municipality. Then we have taken a maximum value of 500 km. for the threshold distance. The profit obtained by chain A from its stores ranges from 10% to 90%. The maximum number of new stores to be located by chain A is 5, which is the number of existing stores of chain B. The parameter δ is taken as $1-\gamma$ in all cases, which means that cannibalism is fully compensated. The following values have been taken as parameters:

- *Threshold value* : $D = 0$, 100, 200, 300, 400 500
- *Unit profit* : $\gamma = 0.1$, 0.2 , 0.3 ,......, 0.8 , 0.9
- *Number of new stores* : $r = 1$, 2 ,...., 5

The following problems have been solved by using the optimizer [27]:

- *Threshold distance (TD) model*: 30 problems
  $r = 1, ..., 5$
  $D = 0$, 100 , 200 , 300 , 400 , 500
- *Side payment (SP) model*: 45 problems
  $r = 1, ..., 5$
  γ =0.1 , 0.2 , ..., 0.9

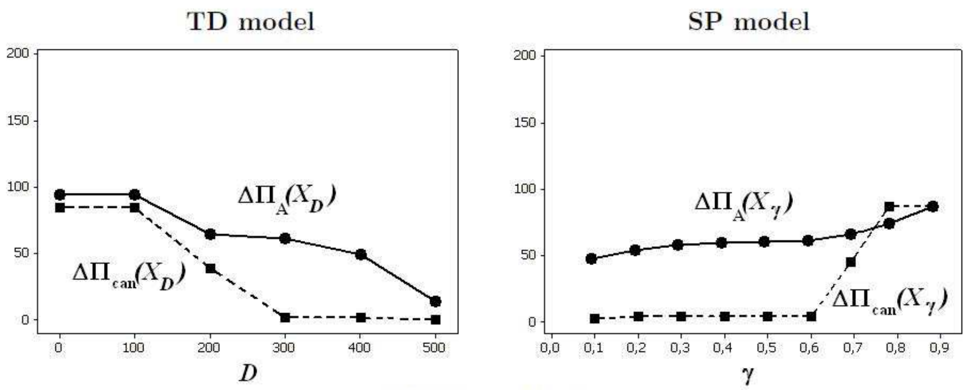
The percentage of profit increase of the owner of chain A and the percentage of cannibalized profit, corresponding to the optimal solutions of the two models, are shown in Figures



$\triangle\Pi_A(X_\gamma) > \triangle\Pi_A(X_D)$

γ ≥	D ≥
0.7	200
0.1	400

Figure 5: Percentages of profit increase and cannibalization for $r = 1$



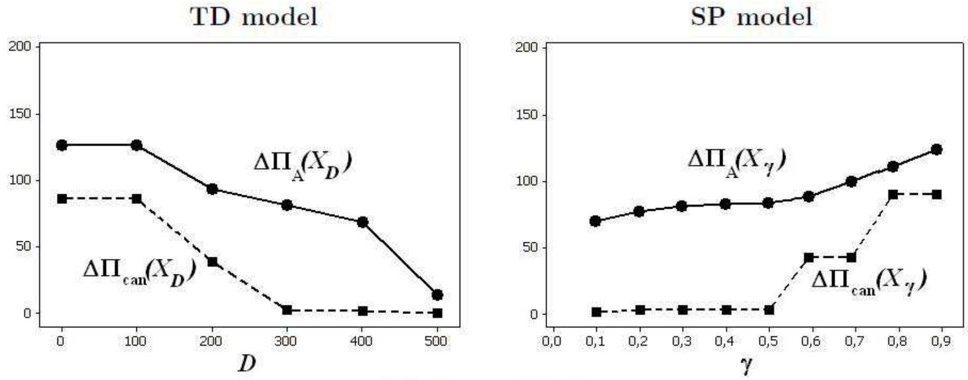
$\triangle\Pi_A(X_\gamma) > \triangle\Pi_A(X_D)$

γ ≥	D ≥
0.7	200
0.2	400
0.1	500

Figure 6: Percentages of profit increase and cannibalization for $r = 2$

5 to 9. The tables in the figures show the combinations of γ and D values for which the percentage of profit increase obtained for the optimal solution of the SP model is greater than the percentage of profit increase obtained for the optimal solution of the TD model.

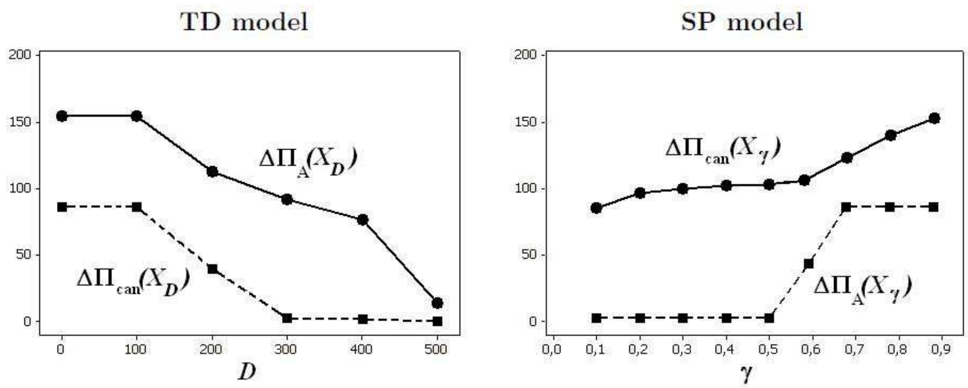
For the TD model, both the percentage of profit increase and the percentage of cannibalized profit are decreasing as long as the threshold distance increases. The rate of decrease in



$\triangle\Pi_A(X_\gamma) > \triangle\Pi_A(X_D)$

γ ≥	D ≥
0.7	200
0.6	300
0.2	400
0.1	500

Figure 7: Percentages of profit increase and cannibalization for $r = 3$

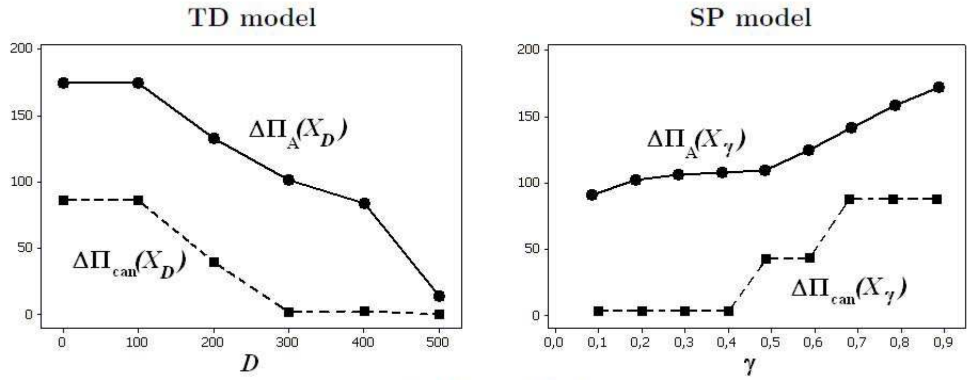


$\triangle\Pi_A(X_\gamma) > \triangle\Pi_A(X_D)$

γ ≥	D ≥
0.7	200
0.2	300
0.1	400

Figure 8: Percentages of profit increase and cannibalization for $r = 4$

both percentages is higher as long as the number of new stores increases. It is observed that the greater the number of stores, the greater the percentage of profit increase. However, the percentage of cannibalized profit remain almost the same when the number of new stores is



$\triangle\Pi_A(X_\gamma) > \triangle\Pi_A(X_D)$

$\gamma \geq$	$D \geq$
0.7	200
0.2	300
0.1	400

Figure 9: Percentages of profit increase and cannibalization for $r = 5$

greater than 1, being almost constant and very low for $D \geq 300$. For the SP model, both the percentage of profit increase and the percentage of cannibalized profit are increasing as long as the amount of money γ obtained by chain A per unit of profit increases. The rate of increase in both percentages is similar as long as the number of new stores increases. The percentage of profit increase grows with the number of new stores while the percentage of cannibalized profit remain almost the same with the exception of a few values of γ in the interval $[0.5, 0.7]$. Notice that the percentage of cannibalized profit is low only for $\gamma \leq 0.4$.

If we compare the results obtained by the two models, the percentage of profit increase with the SP model is greater than the percentage of profit increase with the TD model for a wide range of values of D and γ. The greater value of D, the smaller value of γ is required to hold the previous statement. In such cases, cannibalization with the TD model is low while cannibalization with the SP model is often high. The values of D and γ, for which the percentages of profit increase obtained with the two models are similar, are shown in Table 1. Notice that D is the same for $r = 2, .., 5$, but γ decreases as long as r increases. The run times to generate the optimal locations corresponding to the mentioned values of D and γ are also shown in Table 1. For each value of r, run times with the two models to solve the other proposed problems were similar to the ones shown in Table 1. On average, run times with the TD model were seven times smaller than run times with the SP model. Thus, both models are able to find optimal locations in a few seconds. It is expected that problems of greater size than the ones solved in this paper can be solved in reasonable time and with the TD model is expected to solve problems of greater size than with the SP model.

In all solved problems, the software FICO Xpress Mosel 64 bits v3.10.0 for Linux has been used on a computer with a processor Intel Core i7-6700 3.40 GHz x 8, RAM 8GB, and OS Linux Ubuntu 15.10 64 bits.

r	D	$\Delta\Pi_A(X_D)$	*Run time (sec)*	γ	$\Delta\Pi_A(X_\gamma)$	*Run time (sec)*
1	200	32.20	4.585	0.6	31.01	17.839
2	300	61.17	1.662	0.6	59.91	13.120
3	300	81.24	2.062	0.5	78.98	18.081
4	300	91.87	2.060	0.2	92.41	13.542
5	300	101.09	2.367	0.2	101.94	13.456

Table 1: Similar percentages of profit increase for the two models.

6 Conclusions

We have analyzed the location problem of new stores for an expanding chain which competes with other chains under delivered pricing. The aim of the expanding chain is profit maximization, but taking into account the cannibalization effect. Two approaches have been considered to reduce this effect, one based on a threshold distance and another based on a side payment. If the location space is a transportation network, it is proved that optimal locations can be found in a finite set of points when the threshold distance is used. Then the store location problem is formulated as an integer linear programming model. This model is compared with the corresponding side payment model by using an illustrative example with data from Spanish municipalities. The percentages of profit increase and cannibalized profit corresponding to the optimal locations of the two models are compared. The results obtained for any fixed number of new stores are quite similar. Both percentages are decreasing in D, being the percentage of cannibalization very low for high values of the threshold distance, $D \geq 300$. Both percentages are increasing in γ, being the percentage of cannibalization very low for values of the unit profit not excessively high, $\gamma \leq 0.5$. Therefore, the two models could be used to reduce the cannibalization effect with appropriated values of D and γ.

The choice of the model for locating new stores will depend on each particular case and it will require of a deep study. The main advantage of the SP model is that the cannibalized stores are compensated. As the expanding chain have to pay some money to the cannibalized stores, it could be expected that profit increase with the SP model will be lower than profit increase with the TD model. However, the example shows that the percentage of profit increase obtained with the SP model can be higher than the percentage of profit increase obtained with the TD model. This interesting result occurs for high values of D and γ. In such cases, the SP model is preferred to the TD model. If the percentages of profit increase obtained by the two models are similar, which occurs in the example for some values of D and γ, the SP model is also preferred to the TD model due to the compensation. Otherwise, the TD model could be preferred to the SP model, which happens if the threshold distance is small and the unit profit obtained by the expanding chain from its stores is low.

Both models can be used by managers to estimate profit increase and cannibalization for an expanding chain under delivered pricing competition. The results obtained could be used to make decisions on the type of agreement in franchise systems or in chains where the owner of a new store is different from the owner of the chain. On the other hand, depending on the agreement, the TD model, the SP model, or a mix of these two models, could be of help in decision making on store location. A line for future research is to consider a minimum profit constraint to locate a new store.

Acknowledgments

This research has been supported by the Ministry of Economy and Competitiveness of Spain under the research project MTM2015-70260-P, and the Fundación Séneca (The Agency of Science and Technology of the Region of Murcia) under the research project 19241/PI/14.

References

[1] A. Lancinskas, P. Fernández , B. Pelegrín and J. Zilinskas, (2016), Solution of Discrete Competitive Facility Location Problem for Firm Expansion, *INFORMATICA* 27 , 451-462

[2] Asgari N., Nikbakhsh E., Hill A. and Farahani R.Z. (2016) Supply chain management 1982-2015: a review, *IMA Journal of Management Mathematics* (2016) 27, 353-379.

[3] Bazaraa M.S., Shetty H.D. and Sherali C.M. (1993), *Nonlinear Programming: Theory and Algorithms*. John Wiley & Sons.

[4] Berman O. and Krass D. (2002), Locating multiple competitive facilities: Spatial interaction models with variables expenditures. *Annals of Operations Research* 11, 197-225.

[5] Chandy R.K. and Tellis G. J. (1998), Organizing for radical product innovation: The overlooked role of willingness to cannibalize. *Journal of Marketing Research* 35, 474-87.

[6] Combs J.G., Michael S.C. and Castrogiovanni G.J. (2004), Franchising: A Review and Avenues to Greater Theoretical Diversity. *Journal of Management* 30, 907-931.

[7] Current J.R. and Storbec J.E. (1994), A Multiobjective Approach to Design Franchise Outlet Networks. *Journal of the Operational Research Society* 45, 71-81.

[8] De Beule M., Van den Poel D. and Van de Weghe N. (2015), Assessing the principles of spatial competition between stores within a retail network. *Applied Geography* 62, 125-135.

[9] Drezner T. (2011), Cannibalization in a competitive environment *International Regional Science Review* 34, 306-322.

[10] Fernández J., Pelegrín B., Plastria F. and Tóth B. (2007), Planar location and design of a new facilityy with inner and outer competition: An interval lexicographical-like solution procedure. *Networks and Spatial Economics* 7, 19-44.

[11] Francis R.L., Lowe T.J. and Tamir A. (2002), Demand Point Aggregation for Location Models. In Drezner Z. and Hamacher H. (Eds.), *Facility Location: Application and Theory*. Springer, pp. 207-232.

[12] García M.D., Pelegrín B. and Fernández P. (2011), Location strategy for a firm under competitive delivered prices. *Annals of Regional Science*. 47: 1-23

[13] Ghosh A. and Craig C.S. (1991), FRANSYS: A franchise distribution system location model. *Journal of Retailing* 67, 466-495.

[14] Ghosh A., McLafferty S. and Craig C.S. (1995), Multifacility retail networks, in Z. Drezner ed. *Facility Location: A Survey of Applications and Methods*, Springer 301-330.

[15] Labbe M, Hakimi SL (1991) Market and location equilibrium for two competitors.*Operations Research* 39: 749-756.

[16] P.J. Lederer and J.F. Thisse (1990), Competitive location on networks under delivered pricing. *Operations Research Letters* 9: 147–153.

[17] Lin C., Kilicay-Ergin N.H. and Okudan G.E. (2011), Agent-based Modeling of Dynamic Pricing Scenarios to Optimize Multiple-generation Product Lines with Cannibalization. *Procedia Computer Science* 6, 311-316.

[18] Mwemezi J.J. anf Huang Y. (2011) Optimal Facility Location on Spherical Surfaces: Algorithm and Application. *New York Science Journal* 4, 21-28.

[19] Mazumdar T., Sivakumar K., and Wilemon D. (1996), Launching new products with cannibalization potential: An optimal timing framework. *Journal of Marketing Theory and Practice* 4, 83-94.

[20] Pelegrín B. , Fernández P., García M.D. and Cano S. (2012), On the location of new facilities for chain expansion under delivered pricing *Omega* 40, 149-158.

[21] Pelegrín B., Suárez R. and Cano S.(2012), Isodistant Points in Competitive Network Facility Location, *TOP* 20, 639-660.

[22] B. Pelegrín, P. Fernández, M.D. García (2016) Profit maximization and reduction of the cannibalization effect in chain expansion, *Annals of Operational Research* 246, 57-75.

[23] Pirkul H., Narasimhan S. and De P. (1987), Firm expansion through franchising: A model and solution procedures. *Decision Sciences* 18, 631-645.

[24] Plastria F. (2005), Avoiding cannibalization and/or competitor reaction in planar single facility location. *Journal of Operations Research* 48, 148-157.

[25] Suárez-Vega R., Santos-Peñate D.R. and Dorta-González P. (2011), Location models and GIS tools for retail site location. *Applied Geography* 35 , 12-22.

[26] Suárez-Vega R. and Santos-Peñate D.R. (2014), The use of GIS tools to support decision-making in the expansion of chain stores, *International Journal of Geographical Information*, 28, 553-569.

[27] FICO Xpress Mosel (2009), Fair Isaac Corporation.