Netw Spat Econ (2010) 10:297–320
DOI 10.1007/s11067-009-9127-6

A GIS-Based Optimization Framework for Competitive Multi-Facility Location-Routing Problem

Burcin Bozkaya · Seda Yanik · Selim Balcisoy

Published online: 23 January 2010

Abstract In a dynamic market setting, firms need to quickly respond to shifting demographics and economic conditions. In this paper, we investigate the problem of determining the optimum set of locations for a firm, which operates a chain of facilities under competition. We consider the objective of maximizing profit, defined as gross profit margin minus logistics costs. We propose a location-routing model where revenue is realized according to probabilistic patronization of customers and routing costs are incurred due to vehicles serving the open facilities from a central depot. We propose a hybrid heuristic optimization methodology for solving this model. The optimal locations are searched for by a Genetic Algorithm while an integrated Tabu Search algorithm is employed for solving the underlying vehicle routing problem. The solution approach is tested on a real dataset of a supermarket chain. The results show that the location decisions made by the proposed methodology lead to increased market share and profit margin, while keeping logistics costs virtually unchanged. Finally, we present a GIS-based framework that can be used to store, analyze and visualize all data as well as model solutions in geographic format.

Keywords Competitive facility location · Location-routing · Meta-heuristics · Genetic algorithm · GIS

1 Introduction

Competitive multi-facility location-routing problems arise in contexts where firms operating a chain of facilities such as retail stores, bank branches or other kinds of chain outlets must decide, in the presence of competitors, where to open the next (set

B. Bozkaya (✉) · S. Balcisoy
Sabanci University, Istanbul, Turkey
e-mail: bbozkaya@sabanciuniv.edu

S. Yanik
Faculty of Management, Macka Campus, Istanbul Technical University, Macka 80680, Istanbul, Turkey

of) facilities and/or to move existing facilities to new locations. Many factors influence such decisions, including the available budget and expansion strategy of the firm, the location of the firm's existing facilities and their performance, location of competitors' facilities, the market size, demographics (e.g. population, age, gender, purchasing power) of the demand-generating population, target customer segments, availability and logistical suitability of candidate locations, and existence of other nearby attractions for generating traffic. Normally, in mature economies, the resulting decisions are expected to "last" for a reasonable period of time, in terms of their effectiveness for generating revenue for the firm. However, dynamics of certain industry sectors such as retailing may dictate rather frequent actions, due to demographic shifts, when it comes to opening and closing facilities.

There are two main goals of this paper: one is to extend the body of knowledge for modeling competitive facility location models where logistics issues are also of concern; and the other is to offer a GIS-based framework coupled with an algorithmic solution approach that could potentially lead to an effective decision support system for location analysts. To serve our first goal, we propose an integrated location-routing model that addresses not only the location decisions faced by a firm, but the routing decisions as well. The objective of this model is to maximize total profit, which is the gross profit margin obtained from sales minus facility opening costs and logistics costs.

The location problem our model addresses involves a scenario where the decision maker, presented with a limited budget, must consider and choose from a candidate set one or more new locations to open a facility. The decision maker may also choose to close one or more of the existing facilities, which would effectively result in either closing the facility for good, or transferring/consolidating it to another location. Then we assume that the market will react to the resulting set of open facilities, including those of the competitor, and a revenue realization will occur. Here we consider a Huff-like behavior (Huff 1964) where customer choice is probabilistic, i.e., customers visit open facilities with probabilities proportional to the attractiveness of each facility and inversely proportional to the distance to or from the facility. We consider many different criteria that influence customers' perception of attractiveness such as the size of the store, product variety, ease of access, parking area, and nearby attractions, and combine them in a generic multiplicative function to calculate a single attractiveness score for each facility. The resulting scores are used to distribute total demand to facilities and calculate the market share of the firm as well as its competitor(s).

The second part of our integrated model involves logistics decisions. In particular, we consider the routing problem of serving the open facilities of the chain from a single central warehouse or depot on a daily basis using a variably sized fleet. In essence, we combine the location decisions that are typically considered strategic and long-term, with routing decisions that are considered day-to-day and operational. The main motivation behind this was to address a need in an extremely dynamic retail sector, which we have observed in emerging economies, where facility location decisions are quickly outdated due to demographic as well as geographic shifts and the purchasing behavior of the customer base.

The solution approach and the GIS framework we propose in this paper serve our second goal, which is a DSS-like system that facilitates effective decision making. Given the complexity of the location-routing problems due to its combinatorial nature, such a system must clearly be backed by an optimization routine, which we develop and present in this paper. Our approach is a hybrid heuristic optimization approach where we design and implement a Genetic Algorithm that guides the search for optimal locations, while we resort to a Tabu Search heuristic algorithm for solving the underlying vehicle routing problems and calculating the routing costs. This latter algorithm is commercially available in ESRI ArcGIS 9.3 software as the Network Analyst Extension. We use the ArcGIS 9.3 environment also as a platform for geographic data storage as well as analysis and visualization of location-routing solutions.

As mentioned above, location-routing decisions in a dynamic emerging market can quickly become obsolete due to shifting demographics and economic conditions. We test our solution approach using real data from such an environment, namely a region of the city of Istanbul, Turkey, where supermarket chains fiercely compete for the market share that emanates from a diverse customer base. Results from this case study indicate that our approach can quickly reveal and evaluate the potential for market share growth by simply re-organizing the configuration of open stores to better respond to the demographic distribution and location of competitors in the region. We show that using the proposed methodology, a firm can increase market share over 15% while keeping logistics costs at bay.

The remainder of this paper is organized as follows: we first provide a brief literature review on Huff-like competitive facility location models. In Section 3, we formally describe our mathematical model, which is formulated as a non-linear integer programming model. In Section 4, we describe the heuristic solution approach we have implemented, which is tightly connected to the GIS platform we have used. In Section 5, we describe results of our case study that derives from a real dataset in the city of Istanbul, Turkey. Finally, we present our concluding remarks in Section 6.

2 Literature review

Competitive facility location problems are widely studied in the literature. We find that the very first introductions of the competitive optimal facility location problem are due to Hakimi (1983) and ReVelle (1986), where new facilities must be located on a network space to compete with a number of existing (competitor) facilities, with the assumption that a customer always visits the nearest open facility.

The particular type of problem we study in this paper is also competitive in nature, but it is primarily based on the notion of customers being free to choose, according to their liking, a *number* of facilities to visit, as opposed to customers always visiting the nearest facility. This notion was introduced originally by Reilly (1929), but it was later refined by Huff (1964). Although the models introduced in both studies are not facility location models per se, they lay the very foundation of probabilistic customer patronage, principles of which have been used in later competitive facility location models such as the one presented in this paper.

According to Huff, customers' patronization of different facilities depends on the attractiveness of each facility, and is also inversely proportional to the distance to the facility. This means customers may not always choose the nearest facility; they may sometimes visit a distant facility because it is more attractive. Nakanishi and Cooper (1974) extend this idea one step further and introduce additional attributes for calculating facility attractiveness in the form of a multiplicative interaction model. They also propose a least squares approach for estimating parameters of their model. The resulting concept is the "probabilistic" visit pattern on behalf of customers, which is radically different from what classical competitive facility location models are designed for. Some early studies in the location literature also make use of this idea; Hodgson (1978, 1981) presents noncompetitive location-allocation models that incorporate the spatial interaction nature of the problem and the fact that patrons may sometimes choose facilities that are further away.

The above notion has not gone unnoticed in the competitive facility location literature and many studies have followed up. Several studies including those by Eiselt and Laporte (1989), Drezner (1994, 1998), Drezner and Drezner (1996), and Berman and Krass (1998) have attempted to merge the areas of competitive facility location and probabilistic consumer choice behavior. Berman and Krass (1998), for instance, additionally address a "flow-capturing" aspect, meaning that customers are attracted to facilities not only from fixed locations like their homes, but also while they are en route to a different destination on the network. A rather more comprehensive study is by Serra and Colome (2001), who extend the *p*-median problem and formulate a competitive location model with an objective of maximum demand capture in a network space. They propose a multiplicative competition interaction model in which the proportion of the demand capture is formulated with the distance and attractiveness value. A meta-heuristic solution method, with a greedy randomized adaptive search in the first phase and a tabu search in the second phase, is employed. Okunuki and Okabe (2002) study the problem also on a network where they define probabilistic customer demand along the links of the network. In their study, an exact computational method is presented to find the optimum location of a store in a competitive environment.

Competitive facility location models based on Huff principles are commonly referred to as "gravity-type models" in the literature. Drezner et al. (2002) develop such a gravity-type model with an objective where a threshold was introduced. Their objective is to minimize the deviation from a market share threshold for the facilities. In this study, the main concern is the survival of a facility ensured by a minimum level of revenue or market share. The problem is formulated in a two-dimensional space and simulations with different threshold levels are executed. Another related study is by Drezner and Drezner (2004). The authors study the problem of locating a single new facility in addition to existing facilities, and propose two algorithms. The first is a Weiszfeld algorithm by which a saddle point is reached. The second one is a branch and bound approach used to calculate an effective lower bound on the optimal solution. In another study by Drezner and Drezner (2007), the same authors focus on the solution methods for a gravity-type *p*-median model, using a heuristic approach. The authors define different decaying functions of customer patronization in relation to distance and attractiveness values. They propose tabu search and steepest descent heuristic methods for the problem.

The concept of store attractiveness and its assessment appears in many studies. For instance, in a paper by Drezner and Drezner (2002), empirical attractiveness values obtained through a survey are compared with the attractiveness scores that are inferred from a competitive location model in which real data on sales and buying power are used. The results show a perfect match of the attractiveness values between empirical and theoretic results. In a similar study, Drezner (2006) investigates the same issue for shopping malls. Drezner shows that distance is less of an important factor in shopping mall selection by customers due to high attractiveness values. She reports the most suitable distance power decay factor as 1.27 for locations in a shopping mall, in contrast to the Huff's result of 3 for small stores in the same grocery category. We use the findings in these two studies for setting the parameters of our own distance decay function, which is described in the next section.

Another important modeling element in the competitive facility location literature is the number of players. In this regard, a typical problem is the leader-follower problem. Dasci and Laporte (2005) present profit maximization models that involve the leader and the follower, who are under different competition conditions. They assume that a consumer patronizes only one facility probabilistically according to the distance, service type and the consumer characteristics. Their model also includes a budget constraint and fixed costs. They conclude that the leader has advantages compared to the follower even if it has cost disadvantages. Fernandez et al. (2006) propose a related model where multiple players are involved. These authors study the single facility location problem for maximizing the profit of both the franchisor and the franchisee. They employ probabilistic customer demand and also include a budget constraint. The authors propose two ways of solving the multi-objective nonlinear model, and present the results of their sensitivity analysis on the budget parameter.

In a very recent stream of research, we see studies parallel to the kind of problem we consider in our paper. Aboolian et al. (2007a) use a competitive location model where a set of new facilities to be placed compete with each other as well as with the existing facilities. Their model is also a gravity-based Huff model and the authors take into consideration the market expansion and cannibalization effect of locating new facilities. In another study, Aboolian et al. (2007b) consider a similar setting where the design factors of a facility are included in the model as decision variables. Demand is formulated as a function of total utility served to the customer, which means demand cannibalization and market expansion factors are also incorporated into the model. The authors present optimal solutions for a limited set of design factors, and heuristic solutions for the rest. Taking this one step further, Aboolian et al. (2008) introduce the price factor as well as attractiveness and distance values that influence demand. The overall objective of the proposed model is to maximize the profit with an optimal price and location decision.

As more recent research suggests, Huff-based competitive facility location problems are getting considerable attention, and various attempts are made to introduce additional factors into the model. Our research is along the same lines, and we contribute by introducing various modeling elements to the problem. In particular, we propose a detailed multiplicative function to model facility attractiveness based on attributes of facilities, introduce a customer expenditure function that depends on customer utility as well as income level, and finally allow the possibility of *closing* a facility in addition to opening new one(s).

One aspect of competitive facility location that is rather neglected in all aforementioned literature is the logistics dimension, i.e., evaluation of location decisions from the viewpoint of logistics costs. Based on our experience in the field, we contend that no location decision can or should be isolated from logistics costs for the duration of the investment period under consideration. As such, another contribution of our work is that we introduce transportation costs in our model, which in turn transforms a traditional competitive facility location model into a competitive location-routing model.

Many studies in the literature attempt to model simultaneous location and routing decisions (see Nagy and Salhi 2007 for an extensive review, and many others cited by their work), however we must note that the model we study in this paper is structurally different from classical location-routing problem. The literature deals mainly with locating a new service facility from which delivery routes originate and reach out to customers. In our setting, the facilities we would like to locate are the delivery destinations or "customers" themselves, for the purpose of providing periodic replenishment of goods from a central warehouse. To the best of our knowledge, a location-routing problem with this modeling feature has not yet been considered. In the next section, we present our model and discuss its location and routing features in more detail.

3 Model

In this section, we present a formal description of our mathematical model. We consider existing and new facility locations of the retail chain as well as those of the competitor, and assume that they are patronized by potential customers in a probabilistic fashion. To formulate the mathematical model, let us first introduce our notation. Let I denote the set of customer locations and $i \in I$ indicate individual customer locations, which we consider in our model as aggregate population demand centers. Furthermore, let E, N, and C denote the set of existing, new and competitor facilities, respectively. By *new* locations, we mean the set of candidate locations the firm considers for opening a new facility. A location solution vector X consists of binary decision variables x_j where $j \in E \cup N$ and $x_j=1$ indicates facility j is open, i.e. an existing facility is kept open or a new facility is established in candidate location j. Let $L(X) = \{j \in E \cup N : x_j = 1\}$ denote the set of indices associated with such open facilities.

To formulate the concept of probabilistic facility patronage in our setting, we build upon the basic concepts introduced by Huff (1964) and Reilly (1929), and further extended by other researchers. Here we assume that each population center patronizes an open facility belonging to set E, N, or C proportional to the facility's attractiveness score and inversely proportional to the distance between the demand center and the facility. In other words, each demand center has a "utility" towards an open facility, which is calculated as shown in (1)

$$u_{ij} = \frac{{A_j}^{\alpha}}{{d_{ij}}^{\beta}} \tag{1}$$

Here, u_{ij} denotes this utility value of demand center $i \in I$ towards an open facility $j \in L(X) \cup C$. The constants $\alpha>0$ and $\beta>1$ are used to adjust for the relative importance of the corresponding term, or to account for concepts like facilities being too far out losing their relative importance. In our study, we assume $\alpha=1$ and $\beta=2$.

Once the utility of each demand-facility pair is calculated, the patronage probabilities of a demand center i against all open facilities are formulated as follows:

$$P_{ij}(X) = \frac{u_{ij}}{\sum\limits_{j \in L(X) \cup C} u_{ij}} \tag{2}$$

where $i \in I$ and $j \in L(X) \cup C$. Under this patronization behavior, customers visit facilities according to probabilities P_{ij} and contribute to the overall revenues that are collected by the firm and its competitors. In our model, we denote the total market size by D, of which the firm and its competitors capture a certain share, depending on the configuration of their stores and the demographics of the population. Total market size D is calculated by evaluating revenues (or profit margins to be precise) that can be collected from each demand center. We assume that the total spending by each demand center is a function of the following two factors: purchasing power or income level, and total utility of the demand center. Mathematically, we formulate this concept through a new formulation that we introduce:

$$f_i(I_i, U_i) = a\frac{1 + me^{-I_i/\tau}e^{-U_i\lambda}}{1 + ne^{-I_i/\tau}e^{-U_i\lambda}} \tag{3}$$

Here f_i represents a non-decreasing function that reflects the overall spending, per unit time period such as per month, of demand center i where I_i indicates the income level or purchasing power at center i, and U_i denotes the total utility of center i calculated as the sum of individual utility values u_{ij} of demand center i towards all open facilities $j \in L(X) \cup C$. The remaining parameters a, m, n, τ, λ are used to adjust the behavior of the function in desirable ways, for instance to account for concept of demand elasticity due to different product categories and consumption behaviors.

The spending function we provide above is illustrated in Fig. 1. This function demonstrates the behavior of an S-curve in two dimensions, namely income and utility, meaning that customers are slow at spending first, when their utility and/or purchasing power is considerably low, and then their spending converge to a certain maximum as the need for shopping items of interest saturates. We adjust parameter a to set the maximum spending and the remaining parameters m, n, τ, λ to adjust the way the spending function converges and the relative influence of the income and utility factors.

This type of a non-decreasing mathematical relationship between consumer spending and its deriving factors has been much employed in the competitive facility location as well as retailing and consumer behavior literatures. For instance, Berman and Krass (2002) introduce use of exponential functions in the competitive facility location literature for defining a non-constant relationship between utility and spending, in addition to linear, step-wise and bounded linear functions. Aboolian et al. (2007a, b, 2008) have built upon the same premise to link spending and price, as well as spending and utility. Similar concepts have been

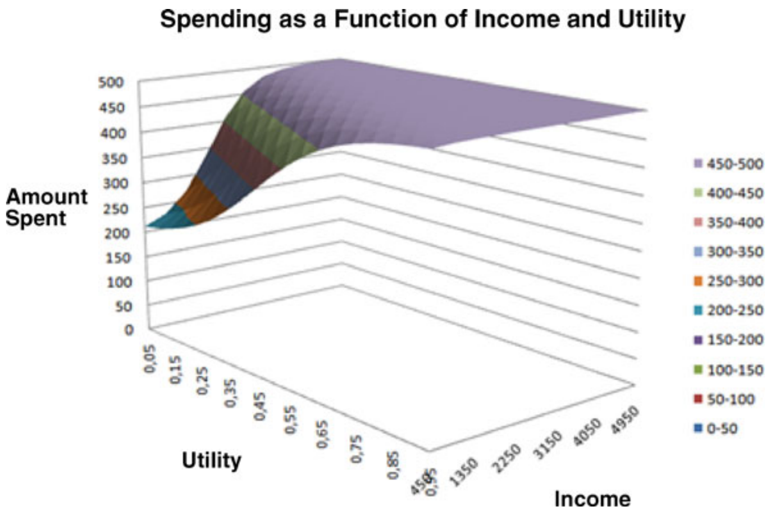

Fig. 1 Spending function for modeling consumer behavior

widely studied in the economics, retailing and consumer behavior literature. For instance, Ghosh and Craig (1991) apply an exponential function to link spending and number of open stores within a certain distance. In economics, income elasticity of expenditure has been explored using Engel elasticity curves and Lorenz curves. There have been both theoretical and empirical studies where the spending patterns of various types of commodities such as food, clothing, furniture, and housing are described using such non-decreasing curves. These studies are rather outside the scope of this text, hence we refer interested readers to the extensive works of Mahalanobis and Kakwani (see Mahalanobis 1960; Kakwani 1978, 1997; Kakwani and Podder 1976) and the empirical study by Abdel-Ghany et al (2002).

The spending function that we propose above generalizes the non-decreasing relationship between consumer spending and the two deriving factors—income and utility—in that it allows for increasing or decreasing levels of elasticity, which may arise with different product categories. Furthermore, this function reduces to different forms as used in some other studies, by setting its parameters a, m, n, τ, λ appropriately. For instance, by setting $n=0$, $a=1$, $m=-1$, and considering only the utility dimension, the spending function in (3) becomes the exponential function used by Berman and Krass (2002), and Aboolian et al. (2007b).

Given this spending function, we calculate the total gross profit margin collected from demand center i (per unit time period) by an open facility $j \in L(X) \cup C$ as follows:

$$r_{ij}(X) = P_{ij}(X) \cdot f_i(I_i, U_i) \cdot t_i \cdot \theta_j \tag{4}$$

where t_i is the population of demand center i, and θ_j is the profit margin of the firm operating facility j applicable to all the revenues collected by the firm. The total profit margin, $R(X)$, collected by all facilities of the chain and the corresponding market share, $\overline{d}$, is then calculated as follows:

$$R(X) = \sum_{i \in I} \sum_{j \in L(X)} r_{ij}(X) \tag{5}$$

$$\overline{d} = \frac{R(X)}{D} \tag{6}$$

$$D = \sum_{i \in I} \sum_{j \in L(X) \cup C} r_{ij}(X) \tag{7}$$

In other words, each potential customer at demand center i spends a portion P_{ij} of his/her, say monthly, spending f_i at facility j, of which a certain portion θ_j is captured by the corresponding firm as profit. What still remains as a building block of our model that needs to be formally described is the calculation of facility attractiveness scores. Similar to the study by Nakanishi and Cooper (1974), we consider various factors that influence the way consumers perceive a facility as attractive, and use a multiplicative model to combine them into a single attractiveness score. For a mathematical representation of this concept, let z_{jl} denote the attractiveness value of facility j with respect to measure l. Without loss of generality, we scale these measure values to the range (0,100]. Furthermore, each measure l has a power weight, denoted by μ_l. Given a set of existing, new and competitor facilities, we calculate $z_l^{\min}$ and $z_l^{\max}$ as the minimum and maximum observed values of measure l, and use these to scale z_{jl} values to the range [0–1]. We obtain the resulting scaled values z'_{jl} as follows:

$$z'_{jl} = \frac{z_{jl} - z_l^{min}}{z_l^{max} - z_l^{min}} \tag{8}$$

We then use these normalized individual scores to calculate the minimum possible attractiveness score A_{min} among all facilities, and the resulting normalized attractiveness score A_j for each facility j.

$$A_{min} = \sqrt[\sum_l \mu_l]{\prod_l \left(z_l^{min}\right)^{\mu_l}} \tag{9}$$

$$A_j = A_{min} + (100 - A_{min}) \left[\left(\sqrt[\sum_l \mu_l]{\prod_l \left(1 + z'_{jl}\right)^{\mu_l}} \right) - 1 \right] \tag{10}$$

In this model, the overall attractiveness score derives from a geometric average of the normalized values under each attractiveness measure l, and is also scaled to the range [A_{min},100]. The exponents μ_l can be set to adjust the relative importance of the factors that influence facility attractiveness. In our study, we consider the following six attractiveness measures, and set their weights μ_l equal to 1:

- Facility size (the larger, the more attractive)
- Ease of access to facility (the easier, the more attractive)
- Size of parking area (the bigger, the better)
- Nearness to other stores and/or attractions (the more attractions in the area, the better)
- Product variety (the more, the better)
- Facility atmosphere and environment (higher scores indicate the better)

Note that, our multiplicative model is similar to the model proposed by Aboolian et al. (2007b). In our formulation, we take additional steps to normalize individual scores and scale them to the range $[A_{min}, 100]$, and hence clearly set a base attractiveness value for all facilities involved.

For the routing aspect of our model, we assume a fleet of maximum $|K|$ vehicles serving each open facility of the chain on a daily basis, where K is the index set for vehicles. Each vehicle has a capacity of b_k units of goods that are to be delivered to an open facility. Vehicles also have time limits $b_k^{'}$ on their routes; that is they must depart and return to the depot, which is represented by index 0, within no more than $b_k^{'}$ units of time. We use the following decision variables to model the route structures and visit sequences:

$y_{jj'k}$: 1 if vehicle k travels from location j to j', $j, j' \in J = E \cup N \cup \{0\}$
w_k: 1 if vehicle k is used in a route, $k \in K$

Finally, our model includes an investment budget constraint that is effective on all facility open/close decisions. This is very much observable in practical settings, as firms implementing an expansion or re-configuration strategy typically allocate certain sums of money for this purpose. In our model, a facility open decision, i.e. choosing one or more of the candidate locations to establish the next set of facilities, consumes the budget, while a facility close decision releases funds that effectively increases the budget by the amount saved per unit time period due to closure. We denote the cost of opening facility at location j, which is annualized over an investment period of five years, by F_j^o, and the annualized savings obtained by closing facility at location j by F_j^c. The available budget is denoted by B. Note that, vehicles, too, have a fixed operational cost F_k^v per day of operation in the planning period. While these costs are also annualized, by multiplying F_k^v by the number of delivery days in a year, they do not consume the "investment" budget associated with opening and closing facilities.

The resulting non-linear integer programming model is presented here for an appreciation of the complete picture:

$$max\, F(X, C) = R(X) - \sum_{j \in N} x_j F_j^o + \sum_{j \in E} \left(1 - x_j\right) F_j^c - \sum_{k \in K} \sum_{j, j' \in J} c_{jj'k} y_{jj'k} - \sum_{k \in K} w_k F_k^v \quad (11)$$

subject to

$$\sum_{j \in N} X_j F_j^o - \sum_{j \in E} \left(1 - X_j\right) F_j^c \le B \quad (12)$$

$$\sum_{j \in J} y_{jj'k} = \sum_{j \in J} y_{j'jk} \quad \forall j' \in E \cup N, \forall k \in K \tag{13}$$

$$\sum_{k \in K} \sum_{j \in J} y_{jj'k} = x_{j'} \quad \forall j' \in E \cup N \tag{14}$$

$$\sum_{k \in K} \sum_{j' \in J} y_{jj'k} = x_j \quad \forall j \in E \cup N \tag{15}$$

$$\sum_{j' \in J} \sum_{j \in E \cup N} a_j y_{jj'k} \le b_k \quad \forall k \in K \tag{16}$$

$$\sum_{j' \in J} \sum_{j \in E \cup N} a'_j y_{jj'k} \le b'_k \quad \forall k \in K \tag{17}$$

$$\sum_{k \in K} \sum_{j \in \{0\}} \sum_{j' \in E \cup N} y_{jj'k} = \sum_{k \in K} w_k \tag{18}$$

$$\sum_{k \in K} \sum_{j \in E \cup N} \sum_{j' \in \{0\}} y_{jj'k} = \sum_{k \in K} w_k \tag{19}$$

$$\sum_{j,j' \in J} y_{jj'k} \le M w_k \quad \forall k \in K \tag{20}$$

$$\sum_{j,j' \in S_h} y_{jj'k} \le |S_h| - 1 \quad \forall k \in K \tag{21}$$

$$x_j \in \{0,1\} \quad \forall j \in E \cup N \tag{22}$$

$$y_{jj'k} \in \{0,1\} \quad \forall j,j' \in E \cup N \cup \{0\}, k \in K \tag{23}$$

$$w_k \in \{0,1\} \quad \forall k \in K \tag{24}$$

In this formulation, we use (13) to ensure that the number of vehicles going into and out of each facility are equal, and (14)–(15) to ensure that vehicles visit only open facilities. Through (16), vehicles are allowed to carry goods only within their capacity limits and through (17) each vehicle has a limit on its daily working hours.

Constraints (18) and (19) enforce idle vehicles not to depart from and return to the depot. Constraint (20) makes sure only those vehicles that travel between facilities and depot are actually used. Finally, we employ a sub-tour elimination constraint in (21) where S_h is the dynamic sub-tour set.

The location-routing model presented above combines location and routing decisions in a generalized fashion and therefore is assumed to have a complexity beyond the limits of exact solution techniques, as far as practical problems are concerned. Indeed, we have tested this assumption, by attempting to solve it for a small problem instance with 5 existing, three candidate, and 15 competitor locations and 100 demand points using CPLEX 11 mixed integer programming solver as well as the constraint programming solver. We were unable to obtain solutions to even this small-sized problem instance; hence we resorted to designing and implementing a heuristic approach. In the next section, we provide details on our solution methodology.

4 Solution methodology and GIS framework

The solution approach we have adopted for solving the competitive multi-facility location-routing problem described in the previous section is a hybrid meta-heuristic approach. We employ two different heuristic algorithms for the location and routing aspects of our problem. We use a Genetic Algorithm (GA) that we have designed and implemented in this study for the location decisions, i.e. which existing facilities to keep open and which candidate locations to use to establish a new facility. Once the set of open facilities are decided, one needs to evaluate the routing cost of the solution and include it in the objective function for further evaluation of the GA population, for which we resort to a Tabu Search algorithm.

Genetic Algorithms are heuristic procedures inspired by biological and evolutionary principles and have been applied to optimization problems extensively. The main idea revolves around the concept of maintaining a pool of solutions for the problem, usually referred to as a population of chromosomes, and continuously improving it through generations of crossovers. Through this process, the genetic material is carried over to offspring and the resulting solutions that are hopefully better in terms of objective function value, or the “fitness” value, are used to replace inferior solutions in the pool. For further details on Genetic Algorithms, the interested reader is referred to the study by Reeves (2003).

Genetic Algorithms have received considerable attention in the facility location literature. Several studies (Hosage and Goodchild 1986; Bianchi and Church 1993; Dibble and Densham 1993; Church and Sorensen 1996; Bozkaya et al. 2002; Estivill-Castro and Torres-Velazquez 1999; Jaramillo et al. 2002; Xiao et al. 2002; Xiao 2006) have attempted to solve variants of p-median, p-center, and other related problems. A recent survey by Xiao et al. (2007) provides a thorough review of these studies from the viewpoint of multi-objective decision making. Our study differs from these past implementations of GA in that a) it uses GA in a hybrid heuristic optimization framework to solve a location-routing problem as opposed to a classical location problem; and b) it uses specific algorithm design elements to address certain model features that are not present in problems attempted in the past. In what follows, we present our methodology, while at the same time discussing how our GA design differs from the existing literature.

4.1 Hybrid metaheuristic solution methodology

In our solution approach, we use the GA concept to govern the general search effort on the optimum set of facility locations to open, i.e. the *X* vector. The general steps of our hybrid methodology are outlined as follows:

1. Construct an initial solution pool of size *n*, and evaluate fitness value of each solution (Tabu Search algorithm is employed here to calculate the routing cost portion of the fitness value). Let F_{min} and F_{max} be the minimum and maximum fitness values in the pool, and X_{min} and X_{max} be the corresponding chromosomes.
2. Repeat for a number of generations, *G*:
   - 2.1. Generate an offspring by executing steps 2.1.1 - 2.1.2:
     - 2.1.1. Select two parent chromosomes P_1 and P_2 from the pool.
     - 2.1.2. Crossover P_1 and P_2 to produce offspring X_o. Let F_o be the fitness value of X_o (use Tabu Search algorithm to calculate routing cost).
   - 2.2. If $F_o > F_{min}$ replace X_{min} with X_o, and set $F_{min} = F_o$.
   - 2.3. If $F_o > F_{max}$, set $F_{max} = F_o$ and $X_{max} = X_o$.
3. STOP and report solution X_{max} as the best solution, and F_{max} as the corresponding objective value.

Let us now describe the details of our Genetic Algorithm:

4.1.1 Encoding

In Genetic Algorithms, encoding of chromosomes plays a critical role in the search for optimal solution(s). An improper encoding scheme not only results in an ineffective search over the solution space, but also increases the computation time. In our case, the formulation of our location-routing model and the fact that we use GA only to decide on location decisions readily lends itself to a binary string representation. The length of each chromosome is $|E|+|N|$, where the first $|E|$ bits represent decisions regarding existing facilities, and the remaining $|N|$ bits represent decisions regarding new facilities. A value of 1 in the binary string means the corresponding facility is open, and 0 means it is closed.

It is important to note here that a binary representation is more suitable for the problem at hand, contrary to other GA implementations in the literature where location indices are used instead of 0-1 digits. The main reason is the fact that in our model, any number of facilities can be open and closed, i.e. there is no requirement to reach a fixed number of open facilities as in models such as the classical *p*-median. It is then natural to use a 0-1 representation to keep the chromosome length fixed. As observed by other researchers (e.g. see Bozkaya et al. 2002), use of binary representation to create solutions with a fixed number of open facilities, say *p*, may result in an ineffective search of the feasible space. In our case, however, this is not an issue due to the nature of our model.

While this encoding is extremely easy to maintain and later use in crossing over parent chromosomes, one needs to make sure the resulting set of open facilities do not violate the budget constraint. In our implementation, we make it a hard rule that a chromosome always respects the budget rule, therefore it is always feasible.

4.1.2 Fitness values

The fitness function we use in our GA is identical to the objective function of the model presented earlier. In other words, the fitness function reflects overall profit, which is total profit margin minus open facility costs plus close facility savings minus fixed and variable routing costs. We use the fitness values generated by this function in assessing the quality of chromosomes and the general quality of the solution pool. Contrary to other studies in the literature (Hosage and Goodchild 1986; Dibble and Densham 1993) no additional infeasibility penalty factor is applied on the fitness value, because the GA always operates on solutions feasible with respect to the budget constraint.

4.1.3 Population size and content

In our implementation, we use a constant-sized population, which is a fairly common feature in the facility location GA literature. That is we take the size of the solution pool *S* as constant. Each chromosome in the population is generated in a random fashion via the following steps: first we initialize the chromosome to the initial state of all facilities, that is, for existing facilities the gene value is 1, and for new facilities the gene value is 0. Then, we iteratively perform three steps: first we generate a list of gene indices whose consideration for modification (i.e. $1 \rightarrow 0$ or $0 \rightarrow 1$) does not violate the budget constraint. Next, we select a random gene index from this list, and set the corresponding gene value. We continue randomly generating indices in the same fashion until all budget is used. This way, we make sure that the resulting solutions always satisfy the budget constraint.

While other studies (e.g. Alp et al. 2003) have implemented additional steps at this stage to ensure equal representation of genes in the population, our implementation does not require such a precaution since the genes represented by binary digits are generated in a completely random fashion.

4.1.4 Parent selection

Two parent chromosomes are needed to produce an offspring. We choose the first parent randomly, and the second parent according to a gene diversity measure that we use. This measure is simply the total number of genes that are different in both parents. For instance, the two parents (1,1,0,0,1,0,1) and (0,0,1,1,0,1,0) have a diversity value of 7, because all of their genes are different. We use this approach to select the second parent from the pool, by simply choosing the one that has the largest diversity value in comparison with the first parent. If there are multiple such chromosomes, we select at random.

Using diversity index in parent selection is a rather uncommon feature of our algorithm. This ensures selection of genetically different parents, increasing the likelihood of generating diverse offspring, thereby increasing the effectiveness of our algorithm.

4.1.5 Generating offspring

Traditional GA crossover operators apply a crossover point to the parent chromosomes and swap portions of the two chromosomes after the crossover point. We did not choose to apply such a strategy since, as a result of such operation, the resulting chromosome may be infeasible with respect to budget, and therefore would either require additional steps to maintain feasibility or penalization of the fitness value. Instead, we take an approach somewhat similar to the one proposed by Alp et al. (2003), which has been shown to perform well, and also allows a better handling of budget violations.

To generate a single offspring from two parents, we proceed as follows: first we copy genes that are identical in both parents to the offspring. For the remaining genes, we randomly choose one of the parents for copying the parent gene value. This gives us the initial offspring which may be infeasible with respect to the budget constraint. If not, the offspring is kept. Otherwise, additional steps are performed to satisfy the budget constraint. For each gene position, we perform the following steps in a greedy remove fashion:

1. If the gene values in the original parent chromosomes are identical, skip this gene position. Such values that appear in both parents are kept and the offspring is not modified.
2. Otherwise, consider reversing the gene value (i.e. $1 \rightarrow 0$ or $0 \rightarrow 1$). If there is a resulting budget improvement, add this gene position to a list of possible alternatives.
3. Select the gene position from the list created in Step 2 that gives the maximum fitness improvement, and modify the gene.
4. If the budget constraint is satisfied, STOP. Otherwise, repeat Steps 1–3 until the budget constraint is satisfied.

As a numerical example, consider the two parents (1,0,0,1,0,0,0) and (1,0,1,1,1,0,1). Let budget consumption values for the individual genes be (−10, −15, −12, 20, 29, 25, 24) and the total available budget be 30 units. Note that negative values indicate budget savings when an existing facility is closed, and positive values indicate the cost of opening a facility. By performing the initialization step above, we first obtain the offspring (1,0,1,1,1,0,0) with a total budget use of −15+20+29=34 units, where the 3rd, 5th and 7th positions are selected at random. Since this value exceeds the available budget, we create a list of gene positions whose modification improves the budget usage. At first, this list is created as {3, 5, 7}. Let's assume position 3 gives us the largest improvement on the fitness value of the offspring. We close existing facility #3 by modifying gene #3 and setting it to 0, update budget use as 34+(−12)=22, and update the feasible gene index list as {5, 7}. Since the resulting budget use is feasible, the corresponding offspring is (1,0,0,1,1,0,1), which we take as the output of the offspring production operation.

4.1.6 Mutation

In Genetic Algorithms, a mutation operator is used to introduce genetic diversity to the population pool. This is done in order to avoid local optimum solutions that may emerge as genetic material in the pool converges to certain schemes. In fact, this is a commonly used feature of GAs, as can be found in the literature. In our implementation, we have chosen not to employ mutation since our initial population as well as the mechanism for parent selection take into account genetic diversity, and hence a mutation adjustment is not likely to provide significant improvements. We have indeed confirmed this observation in our preliminary computations, and decided to implement our GA without mutation.

4.1.7 Replacement

Our replacement strategy takes into consideration the offspring produced from two parents, and uses it to replace the worst member of the pool. We keep track of the worst fitness value and the associated chromosome in the pool, and replace it with the offspring only if the offspring fitness value is better. If not, the offspring is discarded and the algorithm continues to produce a new offspring. While other replacement strategies exist in the literature, such as complete replacement of population or complete replacement except the best solution in the pool, we have chosen the method we have described above, because this method allows continues improvement of the population in terms of average fitness value.

4.1.8 Termination criterion

The GA algorithm generates a single offspring in each generation (or iteration) and terminates when it reaches the maximum number of generations. When it terminates, the best solution in the pool along with its objective function value and market share is reported.

We have tested the above algorithm on a set of randomly generated problem instances that are of equivalent size to the case study problem we solve later. Our goal here was to quickly test the performance of certain parameters of the GA such as population size, and number of iterations, and determine the most suitable settings. After running a number of scenarios, we have chosen to use the following settings for the algorithm: $|S| = 100$, $G = 150$, mutation=0%, and budget tightness is 50%. The latter is governed by an algorithm parameter that ties the average fixed cost $\overline{F}$ of opening a new facility to the number of existing facilities the firm operates. A 100% tightness means that the budget is set equal to $\overline{F}$, leaving room, on average, for only one open facility decision. A 0% tightness means the budget is extremely loose, allowing the firm to open as many new facilities as the number it currently operates.

4.2 GIS framework

Our hybrid heuristic solution methodology requires solving an instance of the vehicle routing problem (VRP) that results from each chromosome created. The VRP cost is then included in the objective function of the solution. The literature for

VRP models and algorithms is vast with many commercial applications, so instead of re-inventing the wheel, we have chosen to use a commercial package. We have selected the ArcGIS 9.3 software platform for this purpose and also for building our GIS-based decision support framework.

ArcGIS is a commercial Geographical Information System (GIS) that has wide use in many broad areas where spatially-enabled data need to be stored, retrieved, analyzed, visualized and even served online. We have used ArcGIS in two ways. First, we use it as a platform to store all problem data in geographic format and visualize them as well as the solutions we obtain through our heuristic approach. Secondly, we use it for running the Network Analyst extension repeatedly to solve the VRP problem. The Network Analyst Extension of ArcGIS allows users to define a network dataset and a vehicle routing problem (among other types of network optimization problems), specify its parameters such as costs, network impedances, network restrictions, and type of output, and solve it to produce multiple routes that originate at one or more depots and visit a set of delivery locations. Figure 2 shows a screenshot of the main Network Analyst problem setup interface.

Network Analyst uses a meta-heuristic algorithm based on Tabu Search for producing VRP solutions. As described in vendor's documentation (ESRI 2009), the VRP solution process first creates an origin-destination matrix of shortest travel costs (or times) between all locations that must be visited by a route. Next, using this matrix, the VRP solver goes through a construction stage by inserting each location one at a time onto the most suitable route. The result of this stage is a feasible initial

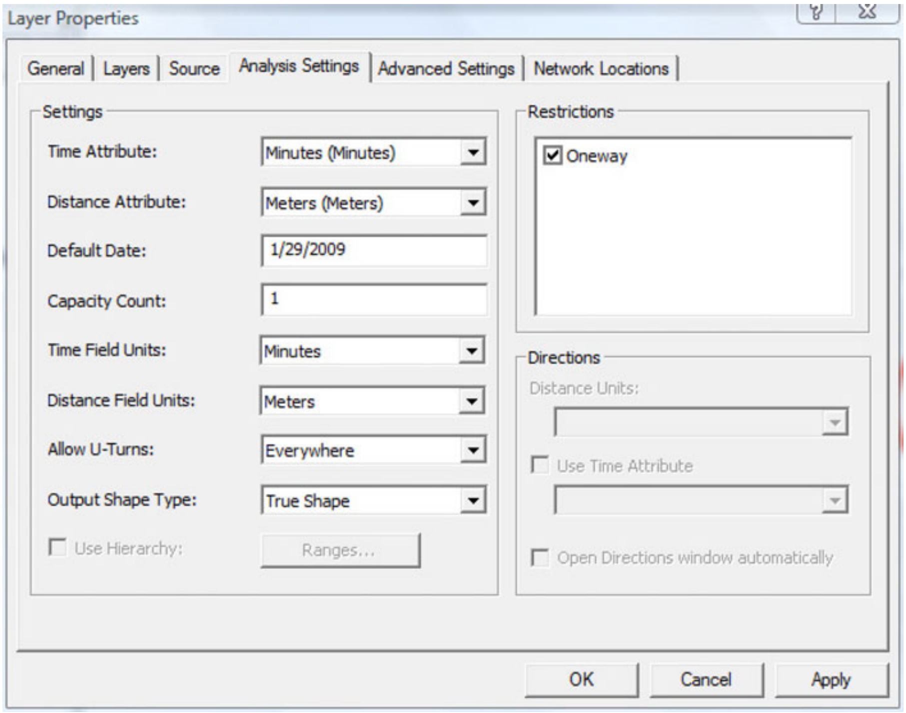

Fig. 2 Network analyst input parameter interface

routing solution. The next step is the improvement stage, where the goal is to search for a better set of routes by performing three types of operations on the initial solution: a) changing the sequence position of a location on a single route, b) moving a single location from its current route to a better route, and c) exchanging two locations between their respective routes. After the algorithm is executed for a pre-set number of iterations and exhausting all possible alternatives to make any further improvement, the algorithm terminates and reports the best solution found. The general execution of the algorithm follows the principles of Tabu Search methodology, where non-improving solutions are accepted along the way, but cycling of solutions are avoided using tabu lists and tabu tenure parameters. (The interested reader is referred to Glover and Laguna (1998) for further details on Tabu Search methodology.) ESRI's solver is proprietary software, therefore further details on the mechanics of this algorithm are unavailable. Nevertheless, we have chosen to use this algorithm because it is a commercially available software module that can generate good solutions fast, which is an essential requirement for our hybrid methodology where instances of the VRP need to be solved repeatedly for members of the GA population during the overall course of the GA execution.

ArcGIS allows customization and access to all of its core objects through a VBA (Visual Basic for Applications) environment. We have used this environment to implement the main loop of the algorithm that uses our GA and the VRP solver algorithm. We have implemented our core Genetic Algorithm in C++, which is called by the main loop until the termination criterion is satisfied. We use the same platform to visualize the solutions generated by our heuristic methodology and present to the analyst. In what follows, we provide further details on the underlying GIS framework in terms of the input layers of geographic data used, the integration of the framework with the GA optimization routines, and visualization as well as reporting of optimization results.

For our optimization routines to work in a GIS environment, several input data layers are stored in ArcGIS's File Geodatabase format in the form of *point* and *line feature class*es.

- Set of existing store locations
- Set of candidate store locations
- Set of competitor store locations
- Set of aggregated demand locations
- Depot location
- Set of vehicles
- Road network dataset

Stored in these input data layers are various data characteristics that are needed for solving the location-routing problem. For existing and candidate store locations, our feature class attribute table includes such attributes as store location, store attractiveness factors and their values, and store shipment amounts for routing purposes. For competitor store locations, the same attributes for existing and candidate stores except shipment amounts are stored. Aggregate demand locations constitute a point representation of demand due to population in a geographic area, where each location point represents a portion of total population as distributed to smaller administrative units known as sub-districts in the area. The depot location layer contains a single point location record that represents the start and end points of routes that are generated by the algorithm. The set of vehicles, defined as a line feature class for later storage of routes built, include such attributes of vehicles as fixed cost of using a vehicle, per kilometer cost of transportation, daily working hours, and total daily shipment capacity. Finally, the road network dataset is the underlying transportation network on which an origin-destination matrix can be calculated and the VRP routing solver can be executed.

In a typical execution of our algorithm, we employ a master VBA routine that integrates the C++ component for GA-related calculations and the ArcGIS-based routing component. This is reflected in Fig. 3. The VBA routine also maintains data transfer between these two components. The VBA routine starts off by reading the geographic data content and internally storing it in memory for sharing between the two components. Any time VBA needs to create a location solution vector which is an offspring chromosome, it calls the C++ component for selecting and crossing over parent chromosomes. The C++ component returns with a location decision vector and its fitness value including all but the routing cost. The VBA routine then calls the VRP subroutine to calculate the routing cost. The VRP subroutine processes the input data passed by the main VBA routine, creates a temporary Network Analyst VRP group layer, runs the Network Analyst VRP solver to create a routing solution, and returns to the main routine with the routing solution as well as its total cost. These steps continue to execute until the pre-defined number of GA iterations are completed, after which the VBA routine reports and visualizes the best solution to the problem. The result of this process is a *line feature class* that stores routes that are built, along with the information on which existing and candidate stores are open and the total objective value of the solution as well as total market share captured by the store chain.

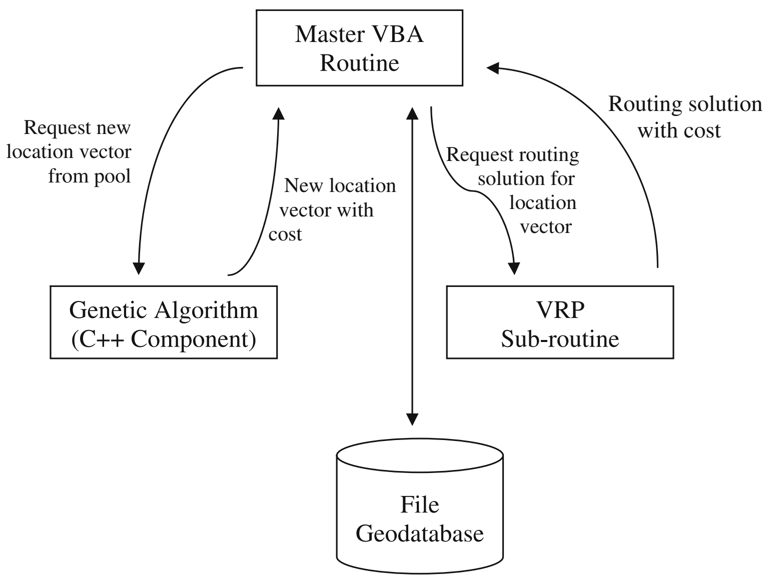

Fig. 3 Main structure of the hybrid heuristic algorithm

5 A case study

In this section, we present the results of our solution approach applied on a real dataset collected from a supermarket store chain in a major metropolitan city. City of Istanbul, the commercial capital of Turkey, is a city with over 12 million inhabitants, according to the results of 2007 census. Several store chains operate in this city, as well as the rest of the country, in direct competition with one another. It is estimated that the annual size of the grocery market in Turkey is over $5 billion, according to a market study done in 2006.

Our region of interest is a section on the eastern side of Istanbul, a densely populated area of approximately 700,000 residents. There are four major retail chains operating in the area, with a total of 33 stores. The geographical area and the distribution of stores are shown in Fig. 4. We consider one of the store chains that has 14 stores as our primary firm, and the remaining three chains with 19 stores as competitors.

There are 28 sub-districts in the study area for which detailed population and income level data are available. Hence, we assume the population in the region is distributed in accordance with sub-district populations, but uniformly within each sub-district. Income level varies across sub-districts as the coastal areas in the southwestern corner of the map have higher income levels, and inner sub-districts have lower. The data we collected also indicate that the overall population is likely to have a monthly grocery spending within the range of $125–325 per household.

To represent this population using the demand centers in our model, we have created a total 168 demand points, on average six per sub-district. Further disaggregation of the population data was not possible due to the effect on computational time.

The current facility configuration with all the demand settings, competitor facility locations, and attractiveness scores for all open facilities indicate that the primary firm of interest can achieve a 41.5% market share. We calculate this figure using the problem data and street-level network travel distances. Here we assume that a profit margin of 5% is applicable overall revenues collected in the region for the primary firm as well as its three competitors, a figure based on our communications with the company officials. Once we have applied our heuristic approach to this dataset with five candidate locations we obtain the results in Table 1 and Fig. 5.

Results in Table 1 indicate that market share and profit increases of at least 15% can potentially be achieved if the facility location configuration can be modified as shown in Fig. 5. This result is promising in a highly competitive market where profit margins are small. With the same amount of logistics resources and a small increase in the total number of open stores, a firm can make higher profits by taking better advantage of the geo-demographics in the region it is operating.

The proposed solution is obtained by solving the competitive location-routing problem on ArcGIS 9.3 platform using the Network Analyst Extension and the Genetic Algorithm we have implemented in C++ as a separate module. With an initial population size of 100 solutions and a total of 150 generations, the algorithm took approximately 20 minutes to complete its execution. Although the computation time may seem excessive for a small problem, it is important to note the algorithm repeatedly solves the VRP problem, where it takes roughly 4 seconds to solve each VRP instance and under 1 second to evaluate the revenue distribution in the region

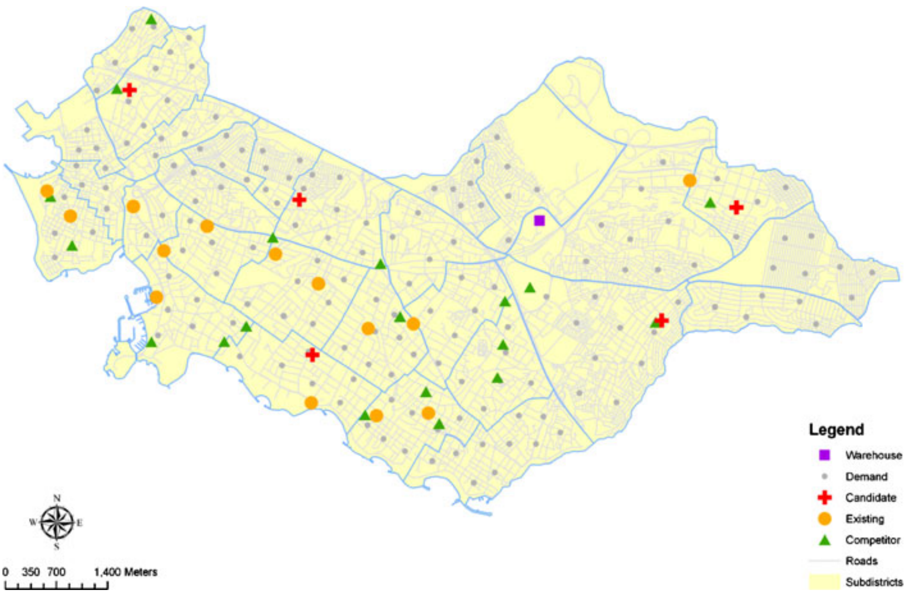

Fig. 4 Istanbul case study problem data

due to probabilistic patronization. This suggests that, in terms of complexity, the VRP problem still remains the most difficult part of the location-routing problem.

6 Conclusion

In this paper, we propose an integrated location-routing model and a hybrid heuristic solution methodology for a competitive multi-facility location problem. The overall model is based on probabilistic patronage of facilities, i.e. customers choosing the facility they will visit proportional to the attractiveness of facilities and inversely proportional to the distance from facility. We also consider logistics costs, as the firm operating a chain of facilities will incur fixed and variable transportation costs for replenishing on a daily basis

Table 1 Results of case study in Istanbul

	Current solution	Proposed solution	% Change
Number of open facilities	14	16	
Number of competitors	19	19	
Number of candidate locations		5	
Number of existing facilities closed		1	
Number of new facilities opened		3	
Total routes used	2	2	
Market share	41.5%	47.9%	+15.4%
Total profit margin	$4,619,018	$5,408,125	+17.1%

Fig. 5 Competitive location-routing solution for Istanbul case study

the facilities that are kept open. The hybrid heuristic algorithm that we propose uses Genetic Algorithm principles to decide which locations to open, and uses a commercially available Tabu Search algorithm to calculate vehicle routing costs associated with the facilities. We use this solution methodology to solve the problem on a real dataset in a considerably dynamic market setting in the city of Istanbul.

One of the contributing features of our approach is that it combines strategic decisions with operational costs within the context of competitive facility location. While this may seem inappropriate in certain settings, we believe the rather dynamic nature of the sector we are studying justifies the use of such a model. In an environment where shifting demographics and changing economic conditions are of concern to decision makers, this model allows analysts to account for more of the relevant decision criteria such as costs, revenues, profit margins, and the possibility of closing a facility no matter how undesirable. The results on a real dataset from the retail sector also indicate that real-world retail systems may have substantial room for improvement. To exploit this potential, however, analysts may need to consider as many relevant criteria as possible. From this point of view, and also considering the modeling elements we have introduced in this paper to the general domain of probabilistic competitive facility location models, we believe our approach is able to capture more realistic scenarios than before.

Several future research directions are possible as a result of this paper. For one, our model can be extended to consider the market share of a firm in a bi-objective manner. In this case, one would expect the firm to simultaneously maximize its gross profit margin and its market share. It is also possible and perhaps even necessary to study the validity of the multiplicative attractiveness model through an empirical study where different attractiveness measures are surveyed. This would be similar to but further extending what is reported by Drezner and Drezner (2002). Finally, there is still an effort needed to turn our proposed solution methodology into a full decision support system in which users can utilize the methodologies presented in this paper in a real application setting.

Acknowledgements This research was supported in part by The Scientific and Technological Research Council of Turkey (TÜBITAK) research grant #105K165. We also would like to extend our thanks to Billur Engin and Emre Koç for their contributions to this study.

References

Abdel-Ghany M, Gehlken A, Silver JL (2002) Estimation of income elasticities from Lorenz concentration curves: application to Canadian micro-data. International Journal of Consumer Studies 26(4):278–285

Aboolian R, Berman O, Krass D (2007a) Competitive facility location model with concave demand. European Journal of Operational Research 181:598–619

Aboolian R, Berman O, Krass D (2007b) Competitive facility location and design problem. European Journal of Operational Research 182:40–62

Aboolian R, Sun Y, Koehler GJ (2008) Optimizing pricing and location decisions for competitive service facilities charging uniform price. Journal of Operational Research Society 59:1506–1519

Alp O, Erkut E, Drezner Z (2003) An efficient genetic algorithm for the p-median problem. Annals of Operations Research 122:21–42

Berman O, Krass D (1998) Flow intercepting spatial interaction model: a new approach to optimal location of competitive facilities. Location Science 6:41–65

Berman O, Krass D (2002) Locating multiple competitive facilities: spatial interaction models with variable expenditures. Annals of Operations Research 111:197–225

Bianchi G, Church R (1993) A non-binary encoded GA for a facility location problem. Working paper, Department of Geography, University of California, Santa Barbara

Bozkaya B, Zhang J, Erkut E (2002) An efficient genetic algorithm for the p-median problem. In: Drezner Z, Hamacher HW (eds) Facility location: Applications and theory. Springer, Berlin, pp 179–205

Church RL, Sorensen P (1996) Integrating normative location models into GIS: problems and prospects with the p-median model. In Longley P, Batty M (eds) Spatial analysis: modelling in a GIS environment. Longman Geoinformation, pp 167–183.

Dasci A, Laporte G (2005) A continuous model for multistore competitive location. Operations Research 53:263–280

Dibble C, Densham PJ (1993) Generating interesting alternatives in GIS and SDSS using genetic algorithms. GIS/LIS ’93 Proceedings 180–189

Drezner T (1994) Optimal continuous location of a retail facility, facility attractiveness, and market share: an interactive model. Journal of Retailing 70(1):49–64

Drezner T (1998) Location of multiple retail facilities with limited budget constraints—in continuous space. Journal of Retailing and Consumer Services 5:173–184

Drezner T (2006) Derived attractiveness of shopping mall. IMA Journal of Management Mathematics 17:349–358

Drezner T, Drezner Z (1996) Competitive facilities: market share and location with random utility. Journal of Regional Science 36:1–15

Drezner T, Drezner Z (2002) Validating the gravity-based competitive location model using inferred attractiveness. Annals of Operations Research 111:227–237

Drezner T, Drezner Z (2004) Finding the optimal solution to the huff based competitive location model. Computational Management Science 1:193–208

Drezner T, Drezner Z (2007) The gravity p-median model. European Journal of Operational Research 179:1239–1251

Drezner T, Drezner Z, Shiode S (2002) A threshold-satisfying competitive location model. Journal of Regional Science 42:287–299

Eiselt HA, Laporte G (1989) The maximum capture problem on a network. Journal of Regional Science 29:433–439

ESRI Inc (2009) Algorithms used by network analyst. In *ArcGIS Desktop 9.3 Help*, web link: http://webhelp.esri.com/arcgisdesktop/9.3/index.cfm?TopicName=Algorithms_used_by_Network_Analyst

Estivill-Castro V, Torres-Velazquez R (1999) Hybrid genetic algorithm for solving the p-median problem. In: Yao X (ed) Proceedings of SEAL '98. The Second Asia-Pacific Conference on Simulated Evolution and Learning. Springer-Verlag, Berlin, pp 18–25

Fernandez J, Toth B, Plastria F, Pelegrin B (2006) Reconciling franchisor and franchisee: a planar biobjective competitive location and design model. In: Seeger A (ed) Recent advances in optimization. Heidelberg, Berlin, pp 375–398

Ghosh A, Craig CS (1991) FRANSYS: a franchise distribution system location model. Journal of Retailing 67(4):466–495

Glover FW, Laguna M (1998) Tabu search. Kluwer Academic, Massachusetts

Hakimi SL (1983) On locating new facilities in a competitive environment. European Journal of Operational Research 12:29–35

Hodgson MJ (1978) Toward more realistic allocation in location-allocation models: an interaction approach. Environment and Planning A 10(11):1273–1285

Hodgson MJ (1981) A location-allocation model maximizing consumers' welfare. Regional Studies 15 (6):493–506

Hosage CM, Goodchild MF (1986) Discrete space location-allocation solutions from genetic algorithms. Annals of Operations Research 6:35–46

Huff DL (1964) Defining and estimating a trading area. Journal of Marketing 28:34–38

Jaramillo JH, Bhadury J, Batta R (2002) On the use of genetic algorithms to solve location problems. Computers & Operations Research 29:761–779

Kakwani N (1978) A new method of estimating Engel elasticities. Journal of Econometrics 8:103–110

Kakwani N (1997) Application of Lorenz curves in economic analysis. Econometrica 45:719–727

Kakwani N, Podder N (1976) Efficient estimation of Lorenz curve and associated inequality measures from grouped observations. Econometrica 44:137–148

Mahalanobis PC (1960) A method of fractile graphical analysis. Econometrica 28:325–351

Nagy G, Salhi S (2007) Location-routing: issues, models and methods. European Journal of Operational Research 177:649–672

Nakanishi M, Cooper LG (1974) Parameter estimate for multiplicative interactive choice model: least squares approach. Journal of Marketing Research 11:303–311

Okunuki K, Okabe A (2002) Solving the huff-based competitive location model on a network with link-based demand. Annals of Operations Research 111:239–252

Reeves C (2003) Genetic alorithms. In: Glover F, Kochenberge GA (eds) Handbook of metaheuristics. Kluwer Academic, Boston, pp 55–82

Reilly WJ (1929) The law of retail gravitation. Knickerbocker, New York

ReVelle C (1986) The maximum capture or 'sphere of influence' location problem: hotelling revisited on a network. Journal of Regional Science 26(2):343–358

Serra DA, Colome R (2001) Consumer choice and optimal locations models: formulations and heuristics. Papers in Regional Science 80:430–464

Xiao N (2006) An evolutionary algorithm for site search problems. Geographical Analysis 38(3):227–247

Xiao N, Bennett DA, Armstrong MP (2002) Using evolutionary algorithms to generate alternatives for multiobjective site search problems. Environment and Planning A 34(4):639–656

Xiao N, Bennett DA, Armstrong MP (2007) Interactive evolutionary approaches to multiobjective spatial decision making: a synthetic review. Computers, Environment and Urban Systems 31:232–252