OR Spectrum (2012) 34:349–370
DOI 10.1007/s00291-011-0280-1

REGULAR ARTICLE

The impact of client choice on preventive healthcare facility network design

Yue Zhang · Oded Berman · Vedat Verter

Published online: 28 December 2011

Abstract In contrast with sick people who need urgent medical attention, the clientele of preventive healthcare have a choice in whether to participate in the programs offered in their region. In order to maximize the total participation to a preventive care program, it is important to incorporate how potential clients choose the facilities to patronize. We study the impact of client choice behavior on the configuration of a preventive care facility network and the resulting level of participation. To this end, we present two alternative models: in the "probabilistic-choice model" a client may patronize each facility with a certain probability, which increases with the attractiveness of the available facilities. In contrast, the "optimal-choice model" stipulates that each client will go to the most attractive facility. In this paper, we assume that the proximity to a facility is the only attractiveness attribute considered by clients. To ensure the quality of care, we impose a bound on the mean waiting time as well as a minimum workload requirement at each open facility. Subject to a total capacity limit, the number of open facilities as well as the location and the capacity (number of servers) of each open facility is the main determinant of the configuration of a facility network.

Y. Zhang
College of Business and Innovation, The University of Toledo,
2801 W. Bancroft St., Toledo, OH 43606, USA
e-mail: yue.zhang@utoledo.edu

O. Berman
Rotman School of Management, University of Toronto,
105 St. George Street, Toronto, ON M5S 3E6, Canada
e-mail: Berman@rotman.utoronto.ca

V. Verter (✉)
Desautels Faculty of Management, McGill University,
1001 Sherbrooke Street West, Montreal, QC H3A 1G5, Canada
e-mail: vedat.verter@mcgill.ca

Both models are formulated as a mixed-integer program. To solve the problems efficiently, we propose a probabilistic search algorithm and a genetic algorithm. Finally, we use the models to analyze the network of mammography centers in Montreal.

Keywords Preventive care · Client choice · Network design · Congestion

1 Introduction

Preventive healthcare involves measures taken to circumvent the development of a disease or for early detection of a condition. It is well established that prevention is more humane and more economical than curing diseases or treating their symptoms. For example, the American Heart Association[1] believes that "basic preventive healthcare services should be an integral part of an equitable, comprehensive healthcare plan, accessible to all" and the decline in the death rates from cardiovascular disease during the past three decades can be attributed to better prevention as well as the improvements in treatment of heart disease and stroke. As another example, from the economical perspective, World Health Organization[2] estimates that people with diabetes generate healthcare costs that are two to three times those without the condition, and in Latin America the costs of lost production due to diabetes are about five times the direct health care costs. In general, preventive services include screenings for certain diseases, immunizations, regular measurements of weight, cholesterol levels, and blood pressure as well as advice about diet, exercise, tobacco, alcohol and drug use, stress, and accident prevention. For example, according to the National Cancer Institute (2011), women should have a Pap test every 3 years to screen for cervical cancer starting at age 21 and a mammogram every 2 years to screen for breast cancer after the age of 50, whereas everyone older than 50 should be tested for colorectal cancer.

Many costly and disabling conditions, including cardiovascular diseases, cancer, diabetes, and chronic respiratory diseases, are linked by common preventable risk factors. World Health Organization (2002) pointed out, however, that the healthcare systems around the globe are primarily based on responding to acute problems and urgent needs of patients, while healthcare providers often fail to seize patient interactions as opportunities to inform patients about health promotion and disease prevention strategies. Nonetheless, there are examples of success that strengthen the motivation to design and improve preventive healthcare programs worldwide. Feachem et al. (2002) reports on the redesign of Kaiser Permanente's primary care clinics to emphasize prevention, patient education, and self-management. Not only the health outcomes have improved for patients with heart disease, asthma, and diabetes, but also the hospital admission rates have declined due to better screening and prevention services. Svitone et al. (2000) reports on a low-cost preventive care program in Ceara, Brazil, that involves monthly home visits by auxiliary health workers, supervised by trained nurses (one nurse to 30 health workers), that has been successful in improving child health status, vaccinations, prenatal care, and cancer screening in women.

1 http://www.americanheart.org.

2 http://www.who.int.

Preventive healthcare is inherently different from healthcare for acute problems. One major difference is that, in contrast with sick people who need urgent medical attention, the clientele of preventive healthcare have a choice in whether to participate in the programs offered in their region and how to choose the facilities they introduce. Empirical evidence suggests that *accessibility* of preventive healthcare facilities plays a key role in determining the level of participation. Zimmerman (1997) found out that the convenience of access to the facility was a very important factor in the clients' decision to have prostate cancer screening. According to McNoe et al. (1996), the main factors for mammography screening non-attendance were practical difficulties and negative attitudes towards the process. The survey by Facione (1999) revealed that the decrease of mammography participation was related to the lack of access. In this paper, we use the proximity to facilities (travel time) as a proxy for accessibility of healthcare facilities.

Assuming that a client always chooses the alternative with the highest attractiveness, many authors used *optimal-choice* in describing the choice behavior. This requires a fully informed and rational set of clients, who patronize their optimal facility at all times. In location theory, for example, it is common to assume that each customer will seek services from the closest open facility. In contrast, under another commonly used assumption, known as *probabilistic-choice*, a client may visit each facility with a certain probability, which increases with the attractiveness of the facility. Probabilistic-choice is a common representation for choice behavior in marketing and econometrics, and it is represented by a variety of spatial interaction models in the literature.

Note that either one of these two assumptions could be realistic depending on a specific circumstance. There are preventive services for which the optimal-choice model is more realistic. For example, when clients are referred to screening facilities by their family physician, it may be safe to assume that they are fully informed about their options and make rational decisions. Some preventive services, however, work on a walk-in basis, e.g., blood tests and vaccinations. The probabilistic-choice model would be more realistic in representing such preventive services.

The objective of this paper was to study the impact of client choice behavior in the preventive healthcare sector. Given that the main focus in allocation of resources is usually on acute care, there are often insufficient resources allocated to preventive care. Consequently, the efficiency of preventive care services is essential. This efficiency can be measured by the participation in the program. In this paper, we focus on the design of a facility network to improve the accessability and thus the participation. This requires a solid understanding of the client choice behavior in reality. To this end, we discuss two models; each one of them assumes a specific client choice behavior and address the question "What is the significance of using the correct client choice model in designing a preventive healthcare facility network?" In other words, what are the negative outcomes if we use the wrong client choice model? We present an analytical framework to determine the number and locations of preventive care facilities as well as the number of servers at each facility so as to maximize the total participation in the program offered by the facility network. We compare the two choice behavior models discussed above. To ensure the quality of care, we impose a bound on the mean waiting time as well as a minimum workload requirement at each open facility.

This is the first paper that presents a probabilistic-choice model for preventive healthcare services. The literature on designing preventive healthcare facility network is sparse. To the best of our knowledge, Verter and Lapierre (2002) and Zhang et al. (2009, 2010) are the only papers that are directly relevant to our work, and both use an optimal-choice model to represent client choice. In their seminal paper, Verter and Lapierre (2002) used distance as a proxy for the accessibility of a facility, whereas Zhang et al. (2009, 2010) stipulated that, in making their facility choice, potential clients focus on the time required to receive preventive care (i.e., traveling plus waiting at the facility). In incorporating probabilistic-choice in the context of preventive care, we apply the multinomial logit model (McFadden 1974) taking into consideration the clients who do not participate in the preventive program.

The remainder of the paper is organized as follows: The next section reviews the related literature on spatial interaction models and their use in service facility network design. Section 3 introduces the probabilistic-choice model and the associated optimal-choice model. A solution procedure based on the location-allocation framework that can be used for both models as well as its computational performance is outlined in Sect. 4. An illustrative example, the network of mammography centers in Montreal, is examined in Sect. 5. The final section presents our concluding remarks. In order to keep the focus of this paper on the modeling aspects and managerial insights, the technical details of the proposed algorithms and our computational experiments are presented in the Appendix.

2 Related literature

The first spatial interaction model that represents the probabilistic-choice behavior is by Huff (1962). His model was developed for analyzing the market share of retail stores, which is formulated as follows:

$$m_{ij} = \frac{U_{ij}}{\sum_{k \in S} U_{ik}}, \tag{1}$$

where m_{ij} is the market share of facility j at population zone i, U_{ij} represents the utility of clients at population zone i patronizing facility j, and S is the set of facilities.

Following Huff's model in which client utility is expressed by a simple gravity formula, a variety of studies in econometrics and marketing focused on developing better representations of the utility function, such as the multiplicative competitive interaction (MCI) model (Nakanishi and Cooper 1974). Both Huff's model and the MCI model were developed based on aggregate flows between population zones and facilities. Another well-known model, the multinomial logit (MNL) model (also called the conditional logit model, McFadden 1974), was originally proposed at the disaggregate (individual) level. Recently, this model was also applied at the aggregate level (Gupta et al. 1996). Denoting by y_{ijl} the lth attractiveness determinant of facility j at population zone i, U_{ij} in the MNL model can be defined as

$$U_{ij} = \exp\left(\sum_{l \in L} \beta_l y_{ijl}\right), \tag{2}$$

where L is the set of attractiveness determinants, and β_l is a parameter that denotes the sensitivity to the corresponding attractiveness determinant and can be estimated empirically. Also note that we assume that β is the same for all j which is reasonable when the problem is studied in the aggregate level (e.g. in the case study presented later, as the clients are older females in the urban area, it is reasonable to assume that they have the same time sensitivity).

In our model we assume that the proximity to the facility is the primary attractiveness determinant. In general, there are other important attractiveness features, such as reputation of the facility, that should be considered. In principle, our models can be generalized to consider such features if data is available.

A number of researchers incorporated the spatial interaction models in designing networks of service facilities. Achabal et al. (1982) might be the first paper to present an MCI-integrated p-median model for selecting multiple locations for a retail chain. Drezner (1994, 1998) considered two MCI-type models for locating new retail facilities to maximize total market share. Berman and Krass (1998) and Okunuki and Okabe (2002) proposed models that integrate Huff's formulation for locating a single or multiple new competitive facilities. Marianov et al. (2008) proposed a facility location problem with congestion by using the MNL model for client allocation, in which travel time and waiting time are considered as the attractiveness determinants.

A significant majority of this literature assumes that all clients would require service and the total market size is fixed. The notable exceptions are Berman and Krass (2002), Aboolian et al. (2007a,b), and Drezner and Drezner (2010). The first three papers incorporate a variable expenditure function in the classical MCI model. They assumed that at each population zone there is a fixed number of clients, who patronize the facilities with frequencies based on the MCI model. The expenditure of clients at a facility is a non-decreasing function of the utility, and the total market for the service expands as new facilities are added. Drezner and Drezner (2010) proposed a new approach to present lost demand in the MCI models by adding a "dummy" facility.

All of the aforementioned studies that address lost demand are based on the MCI model. In contrast, we introduce an MNL model to present it. This approach has not been used in the location models, though it is common in econometrics and marketing. Structurally, it is similar to the one proposed by Drezner and Drezner (2010), i.e., adding an additional term in the denominator of expression (1) that represents the utility of not patronizing any facility. However, the approach based on the MNL model has a more rigorous foundation and also results in fewer parameters to be considered (see Sect. 3.1).

3 The modeling framework

In this section, we first provide the notation and the assumptions that underline our analytical framework. Then, we present the probabilistic-choice model for preventive healthcare facility network design and the analogous optimal-choice model.

Let $G = (N, E)$ be a network with a set of nodes $N(|N| = n)$ and a set of links E. The nodes represent the neighborhoods of population zones, and the links are the main transportation arteries. The fraction of clients residing at zone i is denoted by $h_i, i \in N$. We assume that the number of clients who require service over the entire network is Poisson distributed with a rate of λ per unit of time, and thus from each zone i is also Poisson at a rate $\lambda h_i, i \in N$. We also assume that there is a finite set of potential locations $X \subset N$ for the facilities. Let $S \subset X$ be a set of facilities located.

The travel time from zone i to node j through the shortest path is denoted by t_{ij}. As mentioned earlier, proximity to the facilities (i.e., travel time t_{ij}) is assumed to be the primary attractiveness determinant. We assume that there are $Q_{\max}$ homogeneous servers available, each can provide an exponentially distributed service at a rate of μ services per unit of time, and at least one server must be allocated to each open facility.

Thus, each facility can be modeled as an M/M/c queuing system, where c denotes the number of servers at the facility. The mean waiting time in the system is denoted by $\overline{W}$, and the general formula for this is (Kleinrock 1975)

$$\overline{W} = \frac{C(c, u)}{c} \frac{1}{\mu(1 - \rho)} + \frac{1}{\mu}, \tag{3}$$

where

$$u = \frac{\lambda}{\mu}, \quad \rho = \frac{\lambda}{c\mu}, \quad C(c, u) = \frac{1 - K(u)}{1 - \rho K(u)}, \quad K(u) = \frac{\sum_{l=0}^{c-1} \frac{u^l}{l!}}{\sum_{l=0}^{c} \frac{u^l}{l!}}.$$

The functions $C(c, u)$ and $k(u)$ are just mathematical expressions to simplify (3).

To guarantee a timely service, we assume that the mean waiting time at any facility j denoted by $\overline{W}_j$ cannot exceed a maximum acceptable level denoted by $\overline{W}_{\max}$. We note that other types of service level constraints can be applied in our model as well. In the literature, similar service level constraints have been used in Marianov and Serra (1998, 2002), Wang et al. (2002), and Berman and Drezner (2006). Such service level guarantees are also widely used in practice. For example, the Prime Minister of Canada Stephen Harper has recently announced a new government policy of developing a healthcare guarantee that ensures patients receiving essential medical treatment within clinically acceptable waiting times (Prime Minister of Canada 2007).

Given the number of servers, using (3), it is easy to calculate the maximum number of clients at a facility that satisfies $\overline{W} \leq \overline{W}_{\max}$. Therefore, we perform the following parameter settings to facilitate our model formulations. Suppose that at most K servers can be allocated to each facility. Define $\bar{\lambda}_k, k = 1, 2, \ldots, K$, as the maximum participation rate (such that the queuing system does not explode) at a facility with k servers. Letting $\bar{\lambda}_0 = 0$, we also define $\nabla\bar{\lambda}_k = \bar{\lambda}_k - \bar{\lambda}_{k-1}, k = 1, 2, \ldots, K$, as the incremental value of the maximum participation rate.

To ensure service quality, we assume that facilities cannot be operated unless their participation exceeds a minimum workload requirement denoted by $R_{\min}$. For example, in the Montreal example studied in Sect. 5, the Quebec Ministry of Health made a policy decision to require a minimum of 4,000 mammographies per year for facilities to be accredited.

3.1 The probabilistic-choice model

In this section, we present our MNL model that incorporates the clients' choice of not seeking the preventive care services. Let us start by redefining the quantity m_{ij} as

$$m_{ij} = \frac{U_{ij}}{1 + \sum_{k \in S} U_{ik}} = \frac{e^{\sum_{l \in L} \beta_l y_{ijl}}}{1 + \sum_{k \in S} e^{\sum_{l \in L} \beta_l y_{ikl}}}, \tag{4}$$

and define m_{i0}, the probability (or fraction) of clients at zone i who would not visit any facility as

$$m_{i0} = \frac{1}{1 + \sum_{k \in S} U_{ik}} = \frac{1}{1 + \sum_{k \in S} e^{\sum_{l \in L} \beta_l y_{ikl}}}. \tag{5}$$

Note that the "1" in the denominator of expressions (4) and (5) denotes a normalized term to represent the utility of not visiting any facility (Meyer and Eagle 1982; Earle and Sabirianova 2002). This is consistent with the binary logit model, in which there are only two choices ("yes" or "no").

As mentioned earlier, we assume that travel time t_{ij} is the primary attractiveness determinant, i.e., $|L| = 1$. The problem is to find the optimal set of locations and the number of servers at each open facility, so as to maximize the number of total participants, subject to the constraints on the service level $\overline{W}_{\max}$, the workload requirement $R_{\min}$, and the given number of total available servers $Q_{\max}$.

To formulate the problem as a mathematical program, we define two sets of decision variables:

$$s_{jk} = \begin{cases} 1 & \text{if node } j \text{ has } k \text{ or more servers} \\ 0 & \text{otherwise,} \end{cases}$$

a_{ij}: a continuous auxiliary decision variable denoting the probability (or fraction) of clients at zone i who request the service from facility j.

Note that s_{j1} actually denotes the location decision at node j. Based on the MNL model, a_{ij} can be expressed as

$$a_{ij} = \frac{e^{-\beta t_{ij}} s_{j1}}{1 + \sum_{p \in X} e^{-\beta t_{ip}} s_{p1}} \qquad i \in N \qquad j \in X. \tag{6}$$

This formulation ensures that clients must require service only from open facilities. Expression (6) can be rewritten as

$$a_{ij} + \sum_{p \in X} e^{-\beta t_{ip}} a_{ij} s_{p1} = e^{-\beta t_{ij}} s_{j1} \qquad i \in N \qquad j \in X. \tag{7}$$

Since s_{p1} is a binary variable and a_{ij} is a continuous variable, expression (7) can be linearized as follows by defining z_{ijp} as an artificial continuous variable:

$$a_{ij} + \sum_{p \in X} e^{-\beta t_{ip}} z_{ijp} = e^{-\beta t_{ij}} s_{j1} \qquad i \in N \qquad j \in X \tag{8}$$

$$z_{ijp} \le a_{ij} \qquad i \in N \qquad j, p \in X \tag{9}$$

$$z_{ijp} \le M_1 s_{p1} \qquad i \in N \qquad j, p \in X \tag{10}$$

$$z_{ijp} \ge a_{ij} - M_2(1 - s_{p1}) \qquad i \in N \qquad j, p \in X \tag{11}$$

$$z_{ijp} \ge 0 \qquad i \in N \qquad j, p \in X, \tag{12}$$

where M_1 and M_2 denote two big numbers. For this problem, we set them equal to 1, the upper limit of a_{ij}.

Then, the problem can be formulated as a mixed integer program (MIP):

$$\max \lambda \sum_{i=1}^{n} h_i \sum_{j \in X} a_{ij} \tag{13}$$

s.t. (8) to (12) and

$$\sum_{j \in X} \sum_{k=1}^{K} s_{jk} \le Q_{\max} \tag{14}$$

$$s_{jk+1} \le s_{jk} \qquad j \in X \quad k = 1, 2, \ldots, K-1 \tag{15}$$

$$\lambda \sum_{i=1}^{n} h_i a_{ij} \ge R_{\min} s_{j1} \qquad j \in X \tag{16}$$

$$\lambda \sum_{i=1}^{n} h_i a_{ij} \le \sum_{k=1}^{K} \nabla \bar{\lambda}_k s_{jk} \qquad j \in X \tag{17}$$

$$s_{jk} = 0, 1 \qquad j \in X \quad k = 1, 2, \ldots, K \tag{18}$$

Constraint (14) limits the total number of available servers. Constraints (15) ensure that k servers are already allocated before allocating the $(k+1)$th server to a facility. Constraints (16) stipulate that the number of clients at an open facility must satisfy the minimum workload requirement. Constraints (17) ensure that the number of clients at an open facility with k servers cannot exceed the corresponding maximum level $\bar{\lambda}_k$, so as to limit the expected waiting time to at most $\overline{W}_{\max}$.

3.2 The optimal-choice model

In this section, we present the optimal-choice model. We assume that clients wish to obtain services from the facility with the shortest travel time. Following Zhang et al. (2009), we assume that the fraction of clients at zone i who request service from facility j, denoted by a'_{ij}, is a linear decreasing function of travel time t_{ij}:

$$a'_{ij} = \max\{A_{ij} - \gamma t_{ij}, 0\} \qquad i \in N \quad j \in X, \tag{19}$$

where A_{ij} is the fraction of clients at zone i who would visit facility j when $t_{ij} = 0$, i.e., A_{ij} is the intercept of the participation function, and γ is the slope of the participation function. Note that a'_{ij}, unlike a_{ij}, does not depend on the solution S and thus can be calculated in advance. Also A_{ij} can be determined from survey data. For the Montreal case discussed later, $A_{ij} = 0.95$ which was originally used in Verter and Lapierre (2002) and estimated by the Quebec Ministry of Health.

In addition to s_{jk} defined earlier, we define another set of decision variables for the optimal-choice model:

$$x_{ij} = \begin{cases} 1 & \text{if clients from zone } i \text{ require service from node } j \\ 0 & \text{otherwise.} \end{cases}$$

Thus, the problem can be formulated as another MIP:

$$\max \lambda \sum_{i=1}^{n} h_i \sum_{j \in X} a'_{ij} x_{ij} \tag{20}$$

s.t. (14), (15), and

$$\sum_{j \in X} x_{ij} = 1 \qquad i \in N \tag{21}$$

$$x_{ij} \le s_{j1} \qquad i \in N \quad j \in X \tag{22}$$

$$t_{ij} x_{ij} \le t_{ip} + M_3(1 - s_{p1}) \qquad i \in N \quad j, p \in X \tag{23}$$

$$\lambda \sum_{i=1}^{n} h_i a'_{ij} x_{ij} \ge R_{\min} s_{j1} \qquad j \in X \tag{24}$$

$$\lambda \sum_{i=1}^{n} h_i a'_{ij} x_{ij} \le \sum_{k=1}^{K} \nabla \bar{\lambda}_k s_{jk} \qquad j \in X \tag{25}$$

$$x_{ij}, s_{jk} = 0, 1 \qquad i \in N \quad j \in X \quad k = 1, 2, \ldots, K \tag{26}$$

Constraints (21) ensure that each client zone is served by one facility. Constraints (22) stipulate that clients must require service only from open facilities. Constraints (23), where M_3 denotes a big number, stipulate that clients choose the closest open facility. The constant M_3 can be set equal to the biggest t_{ij} value in the network. Similarly, Constraints (24) and (25) guarantee that each open facility must satisfy both $R_{\min}$ and $\overline{W}_{\max}$.

Both problems are formulated as an MIP, which can be solved directly by standard MIP solvers, such as CPLEX. However, our computational experiments (see Sect. 4) show that, although small- and medium-sized instances (with no more than 40 potential facilities and 100 population zones) may be solved by CPLEX in a few hours, large-sized instances (with more than 40 potential facilities and 200 population zones), such as the Montreal case studied in Sect. 5, cannot be solved to optimality even in days. Therefore, we focus on developing accurate and efficient heuristics.

3.3 A small example

In this section, we use a small-sized example to compare the allocation results of the optimal-choice and probabilistic-choice models.

Suppose there are five open facilities and ten population zones. The travel time matrix in hours and the fractions (h_i) are given in Table 1, and $\lambda = 10$ clients/h. Note that the construction of this example is consistent with that for the computational experiments described in Appendix 2. The parameters used in the optimal-choice model include $A_{ij} = 1.0$ and $\gamma = 1.4$, and the parameter for the probabilistic-choice model is $\beta = 2.0$. Note that we are only interested in the allocation results now, which are independent of μ, $Q_{\max}$, $\overline{W}_{\max}$, and $R_{\min}$.

Tables 2 and 3 show the allocation results of the optimal-choice model and the probabilistic-choice model, respectively. Although the values of the total participation are almost identical, it is clear that the allocations of clients to the facilities are very different.

Note that, in this example, client choice behavior (either optimal-choice or probabilistic-choice) does not have much impact on the number of clients from each zone, whereas its impact on the number of clients patronizing each facility is significant. For instance, in the optimal-choice model, facilities 2 and 3 serve the smallest and largest number of clients, respectively; in contrast, the two facilities interchange their roles in the probabilistic-choice model.

Another difference between the two tables is that the variation in the numbers of clients to the facilities is less in the probabilistic-choice model ([1.074, 1.461] clients/h) than in the optimal-choice model ([0.786, 1.900] clients/h). This is since clients from each zone visit all the facilities in the probabilistic-choice model. Furthermore, this also suggests that the number of facilities that satisfy the minimum workload requirement may be greater in the probabilistic-choice model, and thus total participation

Table 1 The travel time matrix and the fractions

Zone	Facility					
	1	2	3	4	5	Fraction (h)
1	0.405	0.683	0.285	1.008	0.875	0.061
2	0.840	0.425	0.775	0.175	0.408	0.152
3	1.038	0.380	0.630	0.600	0.520	0.071
4	0.843	0.530	0.415	0.435	1.070	0.066
5	1.068	0.433	0.700	0.595	1.038	0.095
6	0.310	0.615	0.940	0.533	0.643	0.022
7	0.688	0.468	0.653	0.475	0.818	0.023
8	0.193	0.278	1.053	0.253	0.625	0.177
9	0.198	0.613	0.153	0.780	0.270	0.160
10	0.945	0.438	0.993	0.698	0.223	0.170

Table 2 The allocation result of the optimal-choice model

Zone	Facility					Total
	1	2	3	4	5	
1	0	0	0.367	0	0	0.367
2	0	0	0	1.148	0	1.148
3	0	0.332	0	0	0	0.332
4	0	0	0.277	0	0	0.277
5	0	0.374	0	0	0	0.374
6	0.124	0	0	0	0	0.124
7	0	0.079	0	0	0	0.079
8	1.292	0	0	0	0	1.292
9	0	0	1.257	0	0	1.257
10	0	0	0	0	1.170	1.170
Total	1.416	0.786	1.900	1.148	1.169	6.419

Table 3 The allocation result of the probabilistic-choice model

Zone	Facility					Total
	1	2	3	4	5	
1	0.105	0.060	0.134	0.032	0.041	0.373
2	0.095	0.219	0.109	0.360	0.226	1.009
3	0.035	0.131	0.080	0.084	0.099	0.430
4	0.049	0.091	0.115	0.110	0.031	0.396
5	0.051	0.180	0.106	0.130	0.054	0.521
6	0.045	0.025	0.013	0.029	0.023	0.136
7	0.023	0.036	0.025	0.036	0.018	0.138
8	0.369	0.311	0.066	0.327	0.155	1.228
9	0.308	0.134	0.337	0.096	0.267	1.142
10	0.099	0.273	0.090	0.162	0.420	1.044
Total	1.180	1.461	1.074	1.367	1.334	6.416

may be greater as well. For the same example, suppose now that $\mu = 3$ clients/h, $Q_{\max} = 3$ servers, $\overline{W}_{\max} = 1$ h, and $R_{\min} = 1.5$ clients/h; when the locations are also decision variables, based on the solution methodology introduced in the next section, one 2-server facility (facility 2) and one single-server facility (facility 3) can be opened in the optimal-choice model, while 3 single-server facilities (facilities 1, 2, and 4) can be opened in the probabilistic-choice model.

4 Solution methods and computational experiments

The two problems we discuss are MIP and can be solved by a commercial software such as CPLEX. As will be discussed later, the problems can take days to be solved for large instances. Therefore, if the decision maker cannot wait that long to obtain the results we present several heuristics.

The solution methods developed here are based on a location-allocation framework and can be applied to solve the problems based on both the optimal-choice and the probabilistic-choice assumptions.

Allocation (Alloc P): Given a set of facility locations, allocate clients to the facilities according to the client choice behavior assumptions (optimal-choice or probabilistic-choice), determine the required capacity at each facility, and check whether the solution is feasible. Since we assume that travel time represents the attractiveness of facilities, expanding capacity at a facility to reduce waiting time will not improve its participation. Therefore, given the set of open facilities S, the optimal capacity that should be allocated to each open facility can be directly obtained by allocating sufficient number of servers to satisfy the service level criterion. The feasibility of a solution depends on whether each open facility satisfies the minimum workload requirement as well as whether the total capacity required is within the limit.

Location (Loc P): Determine the best feasible set of locations. A location heuristic, RPRAE (Repeated-Probabilistic-Remove-Add-Exchange), developed in Zhang et al. (2009) based on a probabilistic search, can be used to solve the location problem. We also developed a Genetic Algorithm (GA, Holland 1975) to solve the location problem.

In this approach, **Alloc P** serves as a sub-routine for **Loc P**. For any set of locations, with **Alloc P** the number of clients at each facility and the objective function value can be determined. See the detailed descriptions of the allocation procedure and the two location heuristic algorithms in Appendix 1.

To examine the computational performance of the solution methods, two experiments have been designed for the two models, respectively. In each experiment, we attempt to compare the performances of our heuristic algorithms and IBM ILOG CPLEX 12.2.0. Our heuristic algorithms are coded in C, and all runs are performed on a workstation with 3.0 GHz Intel Core(TM)2 Quad CPU and 16 GB of RAM. The detailed settings and results of both experiments are presented in Appendix 2.

In brief, the first experiment demonstrates that both heuristic algorithms provide accurate solutions compared with the optimums, as the average deviation is less than 1.5%; moreover, they are much more efficient than the CPLEX solver. This suggests that our heuristic algorithms can solve the instances to near-optimality in a short time.

For the second experiment, CPLEX can only solve the instances of the smallest-sized problem set to optimality within 2 h. We have also tested one relatively large-sized instance with 40 facilities and 100 population zones, and CPLEX cannot solve it to optimality by five days. For a even larger instance, such as 40 facilities and 400 population zones, it is reasonable to expect that it cannot be solved to optimality by CPLEX in a month, as the problem size is increasing exponentially. This demonstrates the need for accurate and efficient heuristics. For each instance of the rest eight problem sets in this experiment, using the best solution obtained from either of the two heuristics as a benchmark, we mainly attempt to compare the performances between the two heuristics. We can observe that the CPU run time proliferates with the number of potential facilities for both heuristics, whereas, it only linearly increases with the number of population zones; for the same problem set, both heuristics typically run for a longer time for the probabilistic-choice model than for the optimal-choice model.

When comparing the two heuristics, both experiments show that, in general, RPRAE exhibits slightly better accuracy than the GA procedure; it also runs for a much longer time than the GA procedure. However, for a few instances, RPRAE provides a very poor solution, implying that the GA procedure is more reliable than RPRAE. Since each heuristic has its own advantage, both are used to investigate the Montreal example in the next section, and the better solution is chosen.

Using the same problem sets, another experiment is conducted to validate the observation described in Sect. 3.3 that the variation in the numbers of clients to the facilities is less prevalent in the probabilistic-choice model than in the optimal-choice model. In this experiment, for each instance of the nine problem sets, setting all the potential facilities open, we calculate the coefficient of variation (CV) of the numbers of clients to the facilities, based on both the optimal-choice and the probabilistic-choice models. We can observe that the average CV for the probabilistic-choice model is consistently smaller. This result validates the difference between the two models in allocation. More importantly, as indicated in Sect. 3.3, the difference in allocation may also lead to different location and capacity decisions, since the latter depend on the allocation result.

5 An illustrative case

We now use the two models to investigate the impact of client choice behavior in an illustrative case, the design of a network of mammography centers in Montreal. This case was previously studied in Verter and Lapierre (2002) and Zhang et al. (2009, 2010). We use the case as the basis for a comparative analysis of the optimal-choice and probabilistic-choice models.

The background was the decision of the Quebec Ministry of Health to subsidize mammogram examinations for women between the ages of 50 and 69. There were 194,475 women in Montreal in this age group and 36 facilities with mammographic equipment. The Ministry made a policy decision to require a minimum of 4,000 mammographies per year for facilities to be accredited. The problem was to determine the facility network so as to maximize total participation.

There are 497 population zones representing the spatial distribution of the potential clients. Based on the assumption of 250 working days per year and 8 working hours per day, the number of potential clients in Montreal per hour $\lambda = 194{,}475/250/8 = 97.24$, and the minimum workload requirement $R_{\min} = 4{,}000/250/8 = 2$ clients/h. It is assumed that the service rate is $\mu = 5$ clients/h the maximum allowed waiting time $\overline{W}_{\max} = 0.5$ h. We set $A_{ij} = 0.95$, $\gamma = 0.37$, and $\beta = 2.25$, so that when all the 36 facilities with a single server are accredited the total participation rate is 63.5% for both models. This is consistent with the previous studies.

To study the network configuration (the facility-server distribution) for both models, we conducted a parametric analysis on the number of total available servers $Q_{\max}$. The results for the two models are shown in Tables 4 and 5, respectively. From the tables, we can find a number of interesting results.

There are a few similarities between the two tables. A major similarity concerns the strategy of capacity pooling versus increasing spatial coverage. Both models show that, when the maximum number of servers $Q_{\max}$ is small, capacity is centralized

Table 4 Parametric analysis on Q_{max} for the optimal-choice model

# of servers	# of open facilities	# of facilities with different number of servers					# of unused servers	Total participation (%)	Ave. participation per server (clients/h)
		1	2	3	4	5			
10	2	0	0	0	0	2	0	46.1	4.479
15	6	0	4	1	1	0	0	60.0	3.887
20	14	9	4	1	0	0	0	62.1	3.017
25	20	15	5	0	0	0	0	63.0	2.451
30	20	14	6	0	0	0	4	63.2	2.362

Table 5 Parametric analysis on Q_{max} for the probabilistic-choice model

# of servers	# of open facilities	# of facilities with different number of servers					# of unused servers	Total participation (%)	Ave. participation per server (clients/h)
		1	2	3	4	5			
10	5	0	5	0	0	0	0	30.6	2.980
15	7	0	7	0	0	0	1	39.0	2.712
20	17	14	3	0	0	0	0	54.2	2.658
25	25	25	0	0	0	0	0	58.2	2.264
30	26	26	0	0	0	0	4	58.9	2.202

at just a few facilities; when there are a relatively large number of servers available, more facilities are accredited, increasing spatial coverage. For instance, when Q_{max} = 10 or 15, both tables show that all the open facilities have two or more servers. However, when Q_{max} = 20 or more, there is only one server required at most of the open facilities. This result is due mainly because, when Q_{max} increases, the average participation rate per server decreases. Thus, the effect of the waiting time guarantee $\overline{W}_{max}$ declines, and more facilities can be accredited to satisfy $\overline{W}_{max}$.

Another interesting finding is that the largest number of servers required in both models is 26. The primary reason for this is that people are not sensitive to waiting time. Consequently, although more servers at the existing facilities would result in even less waiting times, participation would remain the same due to the unchanged travel times; on the other hand, if additional servers are allocated to the closed facilities (i.e., open more facilities), not all the open facilities would reach R_{min}. The same principle also applies to the probabilistic-choice model with $Q_{max} = 15$, where all the seven open facilities have two servers and thus the additional one server is redundant.

There are also several differences between the two tables. Most importantly, the network configurations obtained from the two models are quite different. For example, when $Q_{max} = 10$, the optimal-choice model generates a very special network, which consists of only two giant facilities with 5 servers each, while the probabilistic-choice model results in 5 smaller facilities. Even though both models require at most 26 serv-

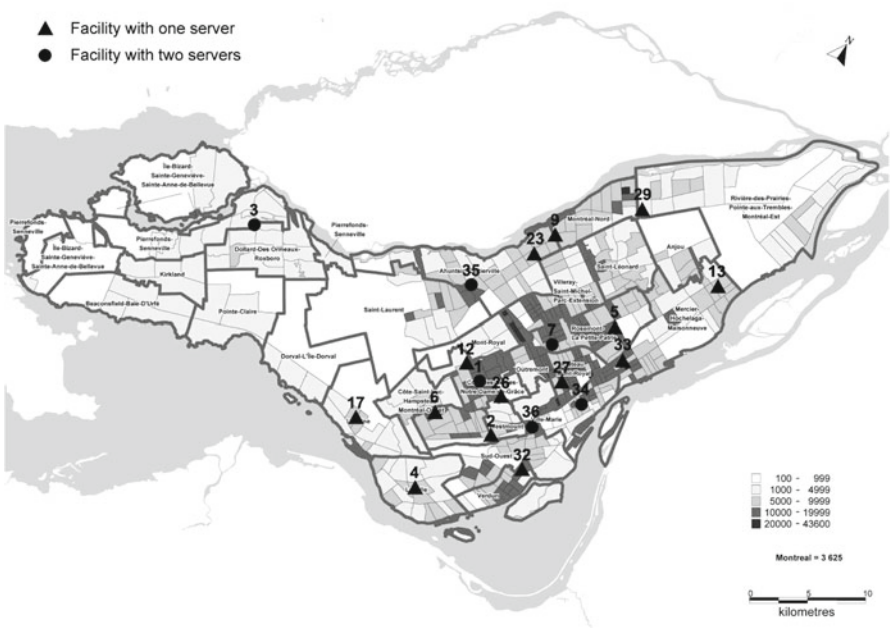

Fig. 1 The solution of the optimal-choice model with 26 servers

ers, the location and capacity decisions are different. Figures 1 and 2 display the two solutions in the population density map of Montreal, respectively. In particular, there are 6 two-server facilities in Fig. 1, while there is only a single server at all the open facilities in Fig. 2.

Based on these comparisons, we can see that centralizing capacity is more necessary for the optimal-choice model in order to satisfy $R_{\min}$ and $\overline{W}_{\max}$ in this case. We conjecture that the main reason for this result is because the probabilistic-choice model typically leads to a lower variation in the numbers of clients to the facilities than the optimal-choice model. For the case of $Q_{\max} = 26$, note that the range of the participation rates at all the open facilities is [2.05, 2.48] clients/h, satisfying both $R_{\min}$ and $\overline{W}_{\max}$. In contrast, if 26 single-server facilities are opened for the optimal-choice model, several facilities cannot reach $R_{\min}$ or exceeds $\overline{W}_{\max}$, and thus capacity has to be centralized at some facilities in high density areas.

In tackling a real-life problem, either the optimal-choice or the probabilistic-choice model will be more appropriate based on the client choice behavior pertaining to the preventive service being offered. It is plausible that data may not be readily available in some cases, and hence the following is a critical question, "What is the significance of using the correct client choice model in designing a preventive care facility network?" In closing this section, we answer this question in the context of the Montreal case.

Suppose *probabilistic-choice* is indeed the real representation of clients' facility choices and there are 26 servers available. Using *the probabilistic-choice model* (the correct model), the highest participation that can be attained is 58.9% (see the last row

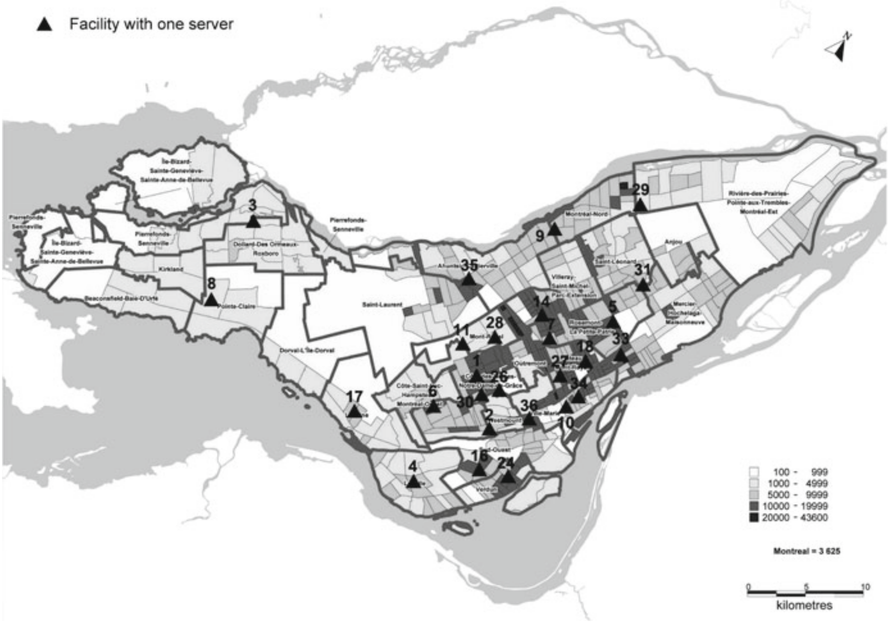

Fig. 2 The solution of the probabilistic-choice model with 26 servers

of Table 5). However, if *the optimal-choice model* (the wrong model) is used, it would lead to a different facility network (i.e., the one shown in the last row of Table 4). Establishing this network in reality, the total participation would be 54.7%. That is, the total participation would decrease by 7.1%, as an ineffective network obtained from the wrong model is established. Similar errors may occur when *optimal-choice* is indeed the real representation but the *the probabilistic-choice model* (the wrong model) is used. Table 4 depicts that, in this case, using the correct model, it is possible to achieve 63.2% participation (see the last row). However, establishing the network obtained from the wrong model (i.e., the one shown in the last row of Table 5), although the total participation would remain the same, 16 facilities would violate either $R_{\min}$ or $\overline{W}_{\max}$, when client allocations are actually made according to optimal-choice. This suggests that client choice behavior has a significant impact on the facility network design. A thorough empirical investigation of the clients to understand their choice behavior is necessary; otherwise, either an ineffective or an infeasible network could be established.

6 Concluding remarks

This paper presents two general and flexible models for designing a service facility network, using the optimal-choice and probabilistic-choice assumptions. Proximity to the facilities is the primary attractiveness determinant. Aiming at maximizing total participation, both models attempt to determine the number of open facilities as well as the location and the capacity of each open facility, subject to a total capacity limit and service level constraints on waiting time and workload.

Both models are formulated as an MIP. To solve the problems efficiently, we proposed two heuristics: one based on probabilistic search and the other is a genetic algorithm. The computational experiments demonstrate that both heuristics can solve the tested random instances to near-optimality in a relatively short time. Using the models, we investigated two illustrative examples, including the network of mammography centers in Montreal, and a few interesting findings are discussed. In particular, the comparison between the two models indicates that client choice behavior may have a significant impact on the location and capacity decisions, and a thorough investigation of this behavior prior to choosing a model is necessary.

Our models can be generalized in several ways. First, although we only use travel time as the primary attractiveness determinant, other attributes, such as facility type, facility reputation, etc., may also be incorporated to the models. Moreover, as indicated earlier, the models and the heuristic algorithms can be revised easily for other types of objective functions, spatial interaction models, service level constraints, or capacity or budget issues.

A direct extension to our probabilistic-choice model that is worth further investigation is to consider waiting time as one of the attractiveness determinants. In this study, we use a service level constraint to capture congestion, and this implies that clients are not sensitive to waiting times as long as it is within a threshold. However, in some cases, waiting time is an important issue for clients to make their choices. Therefore, it would be very interesting to consider the dynamic relationship between client flows to facilities and waiting times at facilities in this probabilistic-choice environment in the future. We note that even though in this paper we use the average waiting time as the measure for waiting time, other measures such as the percentage of customers waiting longer than a given threshold are important and could be used in our models. Nevertheless, some empirical studies in healthcare have shown that the average waiting time is still a significant factor for customer making their choices, for example McGurik and Forell (1984).

Acknowledgments We thank the associate editor and three anonymous referees for their detailed and constructive comments which helped us improve the paper. This research was supported in part by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada (NSERC Grant # 183631).

Appendix 1: Solution methodology

Allocation (Alloc P)

Denoting the feasibility of the solution by V_{fea}, we now provide an allocation procedure for the probabilistic-choice model.

Step 1 For each node $i \in N$ and each facility $j \in S$, calculate a_{ij} according to expressions (6).

Step 2 Calculate the number of clients at each facility $j \in S$, $\lambda_j = \lambda \sum_{i \in N} h_i a_{ij}$, and calculate the objective function value, which is $\sum_{j \in S} \lambda_j$.

Step 3 If $\lambda_j \geq R_{\min} \quad \forall j \in S$, then for each facility $j \in S$, find the smallest number of servers s_j that satisfies $\overline{W}_j \leq \overline{W}_{\max}$. Otherwise, set $V_{\text{fea}} = 0$ (i.e., it is infeasible) and terminate.
Step 4 If $\sum_{j \in S} s_j \leq Q_{\max}$, set $V_{\text{fea}} = 1$ (i.e., it is feasible). Otherwise, set $V_{\text{fea}} = 0$.

Note that for the optimal-choice model, we replace a_{ij} by a'_{ij} and expressions (6) by expressions (19) in Step 1.

Location (Loc P)

Probabilistic search heuristic

This heuristic is composed of the following three basic neighborhood move procedures:

Remove Procedure Starting with all potential facilities open, this procedure removes one facility at a time until feasibility is obtained.
Add Procedure Starting with a feasible facility set, this procedure adds a new facility to the set at each iteration until no addition of facilities can be made while maintaining feasibility.
Exchange Procedure This procedure attempts to improve a given feasible solution by swapping a facility in the current solution with a potential facility that is not currently open and then executing the Add Procedure.

RPRAE runs the Remove, Add, and Exchange Procedures repeatedly for N'_{rep} (a user-defined integer value) times. In each procedure, the heuristic is based on the randomized choice of available alternatives ("alternatives" here can represent facilities to be removed, added, or facility pairs to be swapped). The probability of selecting an alternative is proportional to the change in total demand once this alternative is selected. Refer to Zhang et al. (2009) for a more detailed description. We chose in our experiment $N'_{\text{rep}} = 100$ (this number was used in Zhang et al. (2009) to balance computational time and accuracy).

Genetic algorithm

GA is one of the most successful meta-heuristic for solving combinatorial optimization problems. In GAs, each chromosome represents a solution for the problem, and the quality of a solution is represented by a fitness value. A genetic operator, "crossover", is used to produce new chromosomes from a pair of selected chromosomes, and another operator, "mutation", is used to promote genetic diversity. We refer readers to Reeves (1995) for more details about GAs.

In this paper, a binary coding is used to represent a chromosome. Each chromosome is composed of several binary numbers as genes. Each gene corresponds to the index of a potential facility site, 1 representing an open facility and 0 otherwise. We implement the GA procedure as follows:

Step 0 (Initialization) Randomly generate N_{pop} feasible solutions as a population of chromosomes.

Step 1 (Calculation of the fitness function) For each chromosome in the population, the value of fitness is set to the descending order of its objective function value among the entire population.

Step 2 (Generation of new chromosomes)

Step 2.1 (Parent selection) According to the values of fitness evaluated in Step 1, use the roulette wheel selection method to randomly choose two parent chromosomes from the population.

Step 2.2 (Mutation) For the first parent, randomly choose two genes and interchange their values.

Step 2.3 (Crossover) The chromosomes of the two parents are split into two parts with equal number of genes, and then they are combined across to generate two offspring as new chromosomes.

Step 3 (Replacement) Perform the allocation procedure introduced earlier to calculate the objective function values of the two offspring and to examine their feasibility. For each of the two offspring, if not identical to an existing chromosome, if feasible, and if better than the worst chromosome in the current population in terms of the objective function, this offspring replaces the worst chromosome to keep the population size constant.

Step 4 (Stopping criterion) If the best solution does not change in N_{rep} iterations after the last improvement, then stop and output the best solution from the population. Otherwise, go to Step 1.

Appendix 2: Computational experiments

In the three experiments mentioned in Sect. 4, the number of potential facilities (m) is set to 10, 20, and 40, while the number of population zones (n) is set to 100, 200, and 400. In total, there are nine problem sets, and the number of total available servers $Q_{\max}$ is set to be equal to $(m/2)$. In each problem set, ten instances are generated. For each instance, the demand at each zone (λh_i) is randomly generated in the interval $[0, 2.4(m/n)]$ per hour. The travel times are randomly generated in the interval [0, 1.25] h. The following parameter values were used in the experiments: $\mu = 2.5$ clients/h, $A_{ij} = 1$, $\gamma = 1.4$, and $\beta = 2$, $R_{\min} = 1.2$ clients/h, $\overline{W}_{\max} = 2$ h, and $K = 4$. As indicated earlier, we set M_1 and M_2 equal to 1 for the probabilistic-choice model and set M_3 equal to the biggest t_{ij} value for each instance for the optimal-choice model. Based on preliminary experiments, we set $N'_{\text{rep}} = 100$, $N_{\text{pop}} = 10$, and $N_{\text{rep}} = 500$ for the heuristics.

In the first experiment for the optimal-choice model, we used the optimal solution obtained from CPLEX as a benchmark for each instance. Since it normally takes a very long time for CPLEX to reach optimality for large-sized instances, we only chose the three problem sets with 100 population zones. Moreover, we limited the running time of CPLEX for each instance to 2 h. Table 6 reports the average "Deviation" in total participation and "CPU Time" of the ten instances for each problem set in the first experiment. Note that the results of the last problem sets are based on eight out of the ten instances, for which CPLEX found the optimum solutions within 2 h.

Table 6 The computational performance of the optimal-choice model

# of facilities	# of zones	Deviation (%)		CPU time (s)		
		RPRAE	GA	RPRAE	GA	CPLEX
10	100	0.013	0.178	0.130	0.032	5.768
20	100	1.048	1.446	1.604	0.343	632.371
40	100	0.654	0.964	12.723	2.233	4,517.021

Table 7 The computational performance of the probabilistic-choice model

# of facilities	# of zones	Deviation (%)		CPU time (s)		
		RPRAE	GA	RPRAE	GA	CPLEX
10	100	0.000	0.015	6.290	0.089	271.134
	200	0.000	0.026	6.512	0.141	–
	400	0.000	0.033	7.023	0.240	–
20	100	0.094	0.281	26.742	1.043	–
	200	0.658	0.143	52.399	2.183	–
	400	1.326	0.160	105.433	4.412	–
40	100	0.050	0.181	325.132	16.256	–
	200	0.684	0.219	661.672	33.921	–
	400	0.033	0.178	1,365.761	69.238	–

Table 8 The average CV of the numbers of clients to the facilities based on both models

# of facilities	# of zones	CV	
		Optimal-choice	Probabilistic-choice
10	100	0.364	0.101
	200	0.219	0.060
	400	0.183	0.045
20	100	0.469	0.092
	200	0.352	0.068
	400	0.271	0.050
40	100	0.727	0.099
	200	0.509	0.073
	400	0.381	0.050

In the second experiment for the probabilistic-choice model, we also attempted to use the optimal solution obtained from CPLEX as a benchmark for each instance. However, CPLEX can only find the optimal solution for each instance of the first problem set within 2 h. Hence, for each instance of the rest eight problem sets, we used the best solution obtained from either of the two heuristics as a benchmark, to compare the performance between the two heuristics. Similarly, Table 7 reports the average “Deviation” and “CPU Time” of the ten instances for each problem set in this experiment.

In the third experiment, for each instance of the nine problem sets, setting all the potential facilities open, we calculate the coefficient of variation (CV) of the numbers of clients to the facilities based on both the optimal-choice and the probabilistic-choice models. Table 8 reports the average CV of the ten instances for each problem set based on both models.

References

Aboolian R, Berman O, Krass D (2007a) Competitive facility location model with concave demand. Eur J Oper Res 181:598–619

Aboolian R, Berman O, Krass D (2007b) Competitive facility location and design problem. Eur J Oper Res 182:40–62

Achabal DD, Gorr WL, Mahajan V (1982) MULTILOC: a multiple store location decision model. J Retailing 58(2):5–25

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

Berman O, Krass D (2002) Locating multiple competitive facilities: spatial interaction models with variable expenditures. Ann Oper Res 111(1):197–225

Berman O, Drezner Z (2006) Location of congested capacitated facilities with distance sensitive demand. IIE Trans 38:213–231

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

Drezner T (1998) Location of multiple retail facilities with a limited budget. J Retailing Consumer Services 5:173–184

Drezner T, Drezner Z (2010) Modeling lost demand in competitive facility location. Working paper, California State University-Fullerton

Earle J, Sabirianova K (2002) How late to pay? Understanding wage arrears in Russia. J Labor Econ 20:661–707

Facione NC (1999) Breast cancer screening in relation to access to health services. Oncol Nursing Forum 26:689–696

Feachem GA, Sekhri NK, White KL (2002) Getting more for their dollar: a comparison of the NHS with California’s Kaiser Permanente. Br Med J 324:135–143

Gupta S, Chintagunta PK, Kaul A, Wittink DR (1996) Do household scanner data provide representative inferences from brand choices: a comparison with store data. J Market Res 33:383–398

Holland JH (1975) Adaptation in natural and artificial systems. University of Michigan Press, Ann Arbor

Huff DL (1962) Determination of intra-urban retail trade areas. Real Estate Research Program, UCLA

Kleinrock L (1975) Queueing system I: theory. Wiley, New York

Marianov V, Serra D (1998) Probabilistic maximal covering location-allocation for congested system. J Region Sci 38:401–424

Marianov V, Serra D (2002) Location-allocation of multiple-server service centers with constrained queues or waiting times. Ann Oper Res 111:35–50

Marianov V, Rios M, Icaza MJ (2008) Facility location for market capture when users rank facilities by shorter travel and waiting times. Eur J Oper Res 191:32–44

McFadden D (1974) Conditional logit analysis of quantitative choice behavior. In: Zarembkar P (ed) Frontiers in economics. Academic Press, New York

McGurik MA, Forell FW (1984) Spatial patterns of hospital utilization: the impact of distance and time. Inquiry 21(1):84–95

McNoe B, Richardson AK, Elwood JM (1996) Factors affecting participation in mammography screening. N Z Med J 109:359–362

Meyer RJ, Eagle T (1982) Context-induced parameter instability in a disaggregate-stochastic model of store choice. J Market Res 19:62–71

Nakanishi M, Cooper LG (1974) Parameter estimation for a multiplicative competitive interaction model: least squares approach. J Market Res 11(3):303–311

National Cancer Institute (2011) http://www.cancer.gov/cancertopics/factsheet/Detection

Okunuki K, Okabe A (2002) Solving the huff-based competitive location model on a network with link-based demand. Ann Oper Res 111(1):239–252

Prime Minister of Canada (2007) http://www.pm.gc.ca/eng/media.asp?id=1611

Reeves CR (1995) Modern heuristic techniques for combinatorial problems. McGraw Hill, New York

Svitone EC, Garfield R, Vasconcelos MI, Craveiro VA (2000) Primary health care lessons for the Northeast of Brazil: the Agentes de Saude Program. Pan Am J Public Health 7:293–301

Verter V, Lapierre SD (2002) Location of preventive health care facilities. Ann Oper Res 110:123–132

Wang Q, Batta R, Rump CM (2002) Algorithms for a facility location problem with stochastic customer demand and immobile servers. Ann Oper Res 111:17–34

World Health Organization (2002) Fact sheet No 172. http://www.who.int/mediacentre/factsheets/fs172/en/index.html

Zhang Y, Berman O, Verter V (2009) Incorporating congestion in preventive healthcare facility network design. Eur J Oper Res 198:922–935

Zhang Y, Berman O, Marcotte P, Verter V (2010) A bilevel model for preventive healthcare facility network design with congestion. IIE Trans 42:865–880

Zimmerman S (1997) Factors influencing hispanic participation in prostate cancer screening. Oncol Nursing Forum 24:499–504