

RESEARCH ARTICLE

Robust competitive facility location model with uncertain demand types

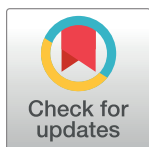
Wuyang Yu *

School of Management, Hangzhou Dianzi University, Hangzhou, Zhejiang Province, China

* yu_wuyang@163.com

Abstract

In competitive settings, firms locate their facilities according to customers' behavior to maximize their market share. A common behavior is consuming from different motivations: one is for convenient demand, and the other is for quality demand. In this behavioral pattern, consumers patronize facilities within convenience for some demands, and patronize high quality facilities beyond convenience range for other demands. This behavior has never been included in competitive facility location problems. Given several other companies' facilities in the market offering similar products or services, we study how a new entrant company can locate facilities based on this customer behavior to maximize its market share. A two-level robust model for the new entrant company is proposed to locate its facilities by taking into account the uncertainty of the types of customers' demands. For medium size problems, we propose an equivalent mixed binary linear programming to obtain exact solutions. For large size problems, an exact algorithm (GCKP-A) for solving the inner-level model is given first by exploring the optimal solution. Then a heuristic algorithm is proposed by imbedding (GCKP-A) and 2-opt strategy into the framework of the improved ranking-based algorithm. The performance of the proposed heuristic algorithm is checked for different size problems. The sensitivity analysis of a quasi-real example shows that: (1) in most cases, the uncertainty between two types of demands does not affect the location scheme; (2) the convenience range, the quality range and the quality threshold play an important role in the market share of the new entrant company.



OPEN ACCESS

Citation: Yu W (2022) Robust competitive facility location model with uncertain demand types. PLoS ONE 17(8): e0273123. <https://doi.org/10.1371/journal.pone.0273123>

Editor: Ashwani Kumar, Sant Longowal Institute of Engineering and Technology, INDIA

Received: May 13, 2022

Accepted: August 2, 2022

Published: August 17, 2022

Copyright: © 2022 Wuyang Yu. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its [Supporting information](#) files.

Funding: The author received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

1 Introduction

Market competition inevitably exists in most business environments. With proper regulation, market competition can promote economic development and technological progress, and provide better services to customers. The competitive facility location problem (CFLP) refers to determining the location of new facilities in a competitive market to maximize the available market share [1, 2]. CFLP arises in a wide variety of applications including locating school, preventive health care facilities, chain stores, park-and-ride facilities, charging stations for electric vehicles, etc. Due to the practicality of CFLP in determining the location of commercial facilities, many models have been proposed to describe various attributes of CFLP from

different aspects, such as customer choice behavior, strategic space, competitive environment, etc [3].

Customer choice behavior refers to the behavior pattern of customers patronizing different facilities that provide similar products or services. Given facilities that can offer similar products or services, customer behavior plays an important role in estimating the market share that each facility can obtain. The binary rule and the proportional rule are the two most common rules used to describe customer choice behavior in the literature [4]. The binary rule assumes that customers always visit the most attractive facilities, it can be traced back to Hotelling's linear market duopoly model [5]. The proportional rule was first proposed by Huff [6], which assumes that customers distribute their purchasing power among all facilities in proportion to the attractiveness of the facility. From the perspective of utilities that customers received from facilities, both the binary rule and the proportional rules are forms of behavior determined by utilities. Due to the diverse behaviors of customers in choosing facilities in reality, scholars have proposed a variety of choice rules to describe different customer behaviors.

A common behavior is consuming from different motivations: one is for convenient demand, and the other is for quality demand. In this behavioral pattern, consumers patronize facilities within convenience for some demands, and patronize high quality facilities beyond convenience range for other demands. For example, people usually go to some nearby places for entertainment on weekends, and for some places with higher attractiveness, such as Disneyland, even if the distance is far, people still choose the right time to visit. This phenomenon shows that there are two types of customer demands: one is for convenience and the other is for quality. In addition, through market research, companies can obtain the value range of the proportion of convenience-type (or quality-type) demands in all customer demands, but it is difficult to obtain the specific value of this proportion in each customer. These are the two motivations of this article: (1) Propose a new rule describing this kind of customer behavior; (2) Study the applicability of robust optimization method in dealing with uncertain demand types. To deal with the first motivation, we considered two different radii, the smaller one is for convenience-type demands, and the larger one is for quality-type demands. We use a threshold to distinguish whether a facility is of high quality or not. Aiming at the second motivation, a two-layer robust model is proposed, in which the inner-layer model is used to solve the impact of uncertain demand types on market share. Applying duality theory to the inner-layer model, after proper linearization, the two-layer robust model can be transformed into a mixed integer linear programming. Therefore, small and medium problems can be solved directly by optimization software. For large-scale problems, we first prove a theorem for the optimal solution of the generalized continuous Knapsack problem. Then we propose an exact algorithm (GCKP-A) to solve the inner-layer model. Then a heuristic algorithm is proposed by embedding (GCKP-A) and 2-opt strategy into the framework of the improved ranking-based algorithm. The performance of the proposed heuristic algorithm is checked on 40 benchmark instances. The sensitivity analysis results of a quasi-real example show the impact of different parameters on the location of competitive facilities.

The paper is structured as follows. Section 2 presents a brief literature review of competitive facility location problems. Section 3 proposes a new customer choice rule to describe the customer behavior presented in section 1. Section 4 establishes the robust competitive facility location model. Section 5 gives the exact algorithm for the inner-layer problem, and proposes a heuristic algorithm by embedding the inner algorithm into the framework of the ranking-based procedure. Section 6 reports the computational results comparing the performance of our approach with other approaches in the literature. Finally, Section 7 concludes.

2 Literature review

Depending on how competitors respond to each other's decisions, the main research streams in the competitive facility location problems can be divided into three categories [7]:

(1) Static competition: In this case, a new firm enters a market and provides similar products or services as existing competitors. The basic assumption is the rivals take no action against the newcomer. For example, Drezner et al. [8] extended the proportional rule by assuming randomly distributed facility attractiveness, they proposed two effective methods to solve the competitive facility location problem to analyze "effective" attractiveness. Fernández et al. [9] studied a variant of the proportional rule, assuming that customers only patronize those facilities that they find attractive to be greater than or equal to a threshold. Marianov et al. [10] introduced the customer choice behavior called comparison-shopping in competitive facility location problem. Comparison-shopping means that customers visit multiple stores selling different products, make comparisons before making a purchase decision. By assuming the utility of a customer consists of two parts: a measurable utility value and its non-observable part, Ljubić and Moreno [11] adopted the multinomial logit model to estimate the captured customer demands and proposed a method to solve the model by combining two types of cutting planes. Ahmadi and Ghezavati [12] used two sustainability measures, flexibility and productivity, to develop an attractiveness function for each facility. They proposed a chance constraint to control the dissatisfaction in the waiting time for services according to the Jackson Markov network. Mahmoodjanloo et al. [13] studies the multimodal competitive hub location pricing problem for a new company planning to enter a market where an existing competitor operates its hub-and-spoke network. They propose a scatter search algorithm based on a nested approach to solve the model. Mai and Lodi [14] investigated the maximum capture facility location problem by assuming that customers choose facilities according to random utilities. Taking advantage of the convexity and separable structure of the objective function, they proposed an enhanced outer approximation algorithm. Aboolian et al. [15] developed a generalized framework of the competitive facility location and design problem, this generalized facility location and design problem include many classic location models as special cases. Fernández et al. [16] used a Pareto-Huff customer choice rule in the competitive facility location problem, and proposed a heuristic procedure to obtain the best approximate solutions.

(2) Dynamic competition: This type of competition is characterized by the assumption that competitors can respond to new entrants. In general, the strategic decisions of competitors are different to change due to high costs, and usually only the operational strategies of competitors are considered. Dynamic competition is often modeled as a Nash equilibrium problem, which is solved using differential systems. For example, Fahimi et al. [17] considered the competitive supply chain network design problem in which two competitors enter the market at the same time. They derived the equilibrium condition and established the finite-dimensional variational inequality formulation, then proposed an algorithm to solve the problem. Zhou et al. [18] proposed a game-theoretical model to study the price competition between two recyclers of electrical and electronic equipment waste. Cai et al. [19] studied the price and warranty competition between two risk-averse retailers by assuming that the retailers' risk-aversion levels are private information. Shalouhi et al. [20] studied Nash and Stackelberg models of competition problems in two pharmaceutical supply chains in which both exclusive retailer and manufacturer are included. Sazvar et al. [21] developed a scenario-based multi-objective programming model to design a sustainable closed-loop pharmaceutical supply chain, which simultaneously considers competitive market and demand uncertainties. Yu and Khan [22] employed stochastic programming and fuzzy mathematical programming to establish a multi-

objective uncertain equilibrium model for the green supply chain network, and they integrated several different methods to solve the model.

(3) Competition with foresight: The newcomer makes logical decisions by forecasting the possible reactions of the competitor after its entry. Qi et al. [23] considered a variant of the proportional rule that customers only patronize facilities within a range they feel is convenient. They proposed a hybrid tabu search algorithm to solve the corresponding competitive facility location problem. Yu [24] studied the leader-follower competitive facility location problem with the partially proportional rule, this rule assumes that a customer chooses the most attractive firm at first then patronizes all facilities of this firm according to proportional rule. Santos-Peñate et al. [25] considered the competitive facility location problem with the binary and S-shaped customer choice rule, they proposed a metaheuristic procedure combining the kernel search algorithm with linear programming for the follower problem. Saha et al. [26] used a multinomial logit model to represent customers' preferences, and proposed a non-linear integer programming model for the joint facility location and inventory problem with partial-disruption risk. Yu and Khan [27] built an evolutionary game model based on the relationship between agricultural product suppliers and urban residents in the financing system. Lin et al. [28] investigates a bilevel competitive facility location problem to maximize expected revenue by considering a discrete choice rule. Latifi et al. [29] studied the leader-follower competitive facility location problem in a closed-loop supply chain, where customer behavior is the Huff gravity-based rule. By replacing the inner level program with its corresponding Karush-Kuhn-Tucker conditions, they proposed an improved branch-and-refine algorithm to solve this problem.

The above literatures mentioned above combine different customer choice rules to determine the market share a facility can obtain, but none of the rules are suitable for describing the customer behavior described in the introduction of this paper. Therefore, in order to better describe this kind of customer behavior, we need to propose a new customer choice rule, and then study the competitive facility location problem on this basis, so as to provide a useful reference for decision makers.

3 New customer choice rule

In this paper, we assume that every customer has two different types of demands. One is the convenience-type demands, which requires the facilities that meet this type of demand are within the convenience range of customers. The other is the quality-type demands, which requires the facilities that meet this type of demand to have sufficient quality and its range is larger than the convenience range. That is, we consider two patronizing radii, a smaller radius (R_c) for convenience-type demands and a larger radius (R_q) for quality-type demands. The latter can only be used for facilities whose quality exceeds a given threshold (δ). From the perspective of customer, the patronage pattern of customers is shown in Fig 1.

It can be seen from Fig 1 that customers only patronize facilities within the range of R_c to meet their convenience-type demands. On the other hand, in order to meet their quality-type demands, customers patronize facilities whose quality exceeds the quality threshold δ and falls within the range of R_q , where $R_q > R_c$. Note that, no matter the demands are convenience-type or quality-type, the customer behavior follows the proportional rule, in other words, his demands are split by these facilities that satisfy corresponding conditions. If a customer exceeds the R_c range of all facilities, his convenience-type demands cannot be met. Similarly, if a customer exceeds the R_q range of all facilities whose quality exceeds δ , his quality-type demands cannot be met. From the perspective of facility, the patronage pattern of customers is shown in Fig 2. It can be seen from Fig 2 that for a facility whose quality exceeds the threshold

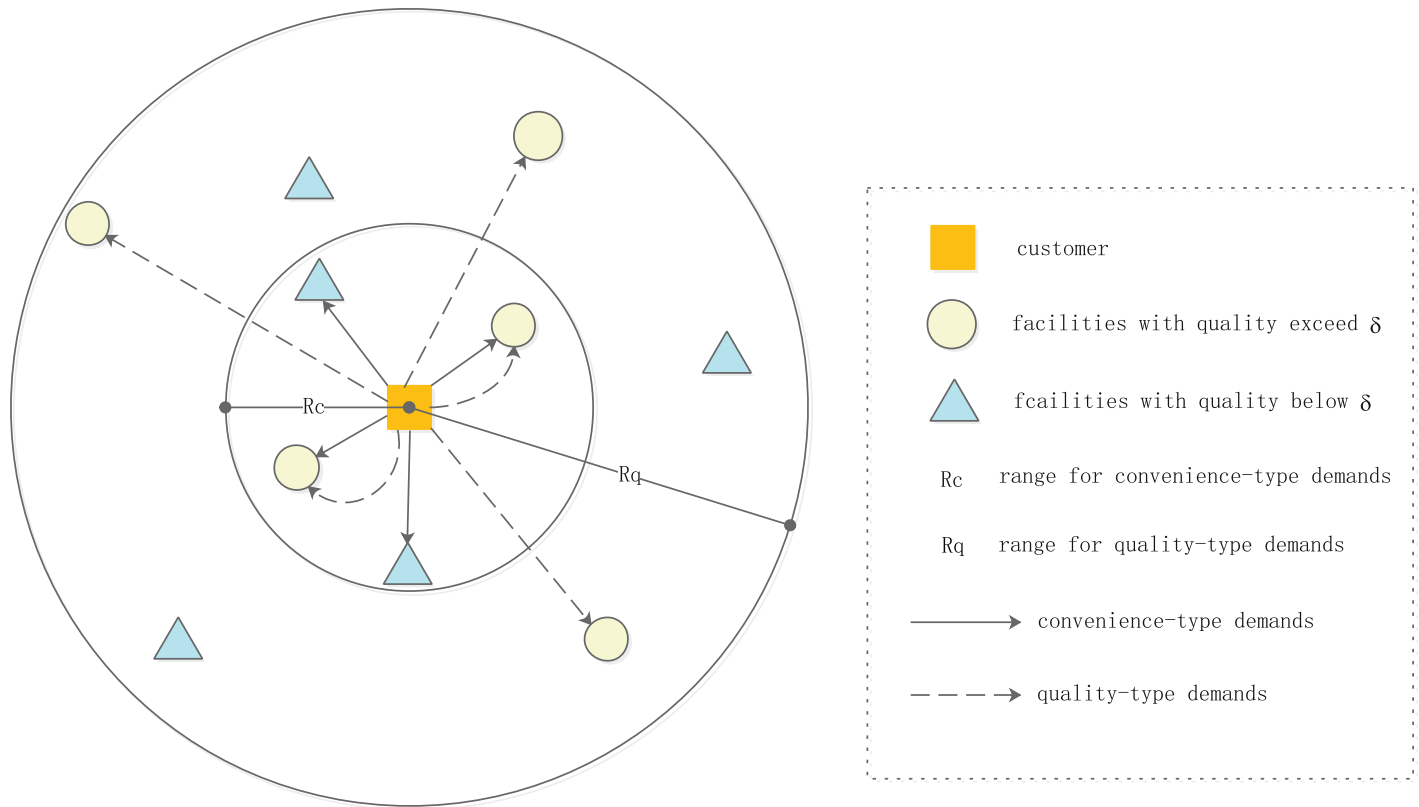


Fig 1. Patronage pattern from the perspective of customer.

<https://doi.org/10.1371/journal.pone.0273123.g001>

δ , it can meet all types of demands within the range of R_c . In addition, it can also meet the quality-type demands between the range of R_c and R_q . However, for a facility whose quality is lower than the threshold δ , it can only meet the convenience-type demands within the range of R_c .

4 Robust competitive facility location model

4.1 The original two-layer model

An entering company hopes to compete for market share of a region by opening several new facilities, and other companies have already established similar facilities. Assumed that the demands of each customer can be divided into two types: the convenience-type demands and the quality-type demands. The total demands of each customer are fixed and known. Through preliminary market research, companies can obtain the value range of the proportion of convenience-type (or quality-type) demands in all customer demands, but it is difficult to obtain the specific value of this proportion in each customer. Due to a limited budget, the new company needs to determine the location of all new facilities from a set of potential locations in order to obtain the largest market share.

The following notations are used:

I : Set of demand points (customers), indexed by i ;

J_E : Set of locations that have established facilities, indexed by j ;

J_N : Set of potential locations for new facilities, indexed by j ;

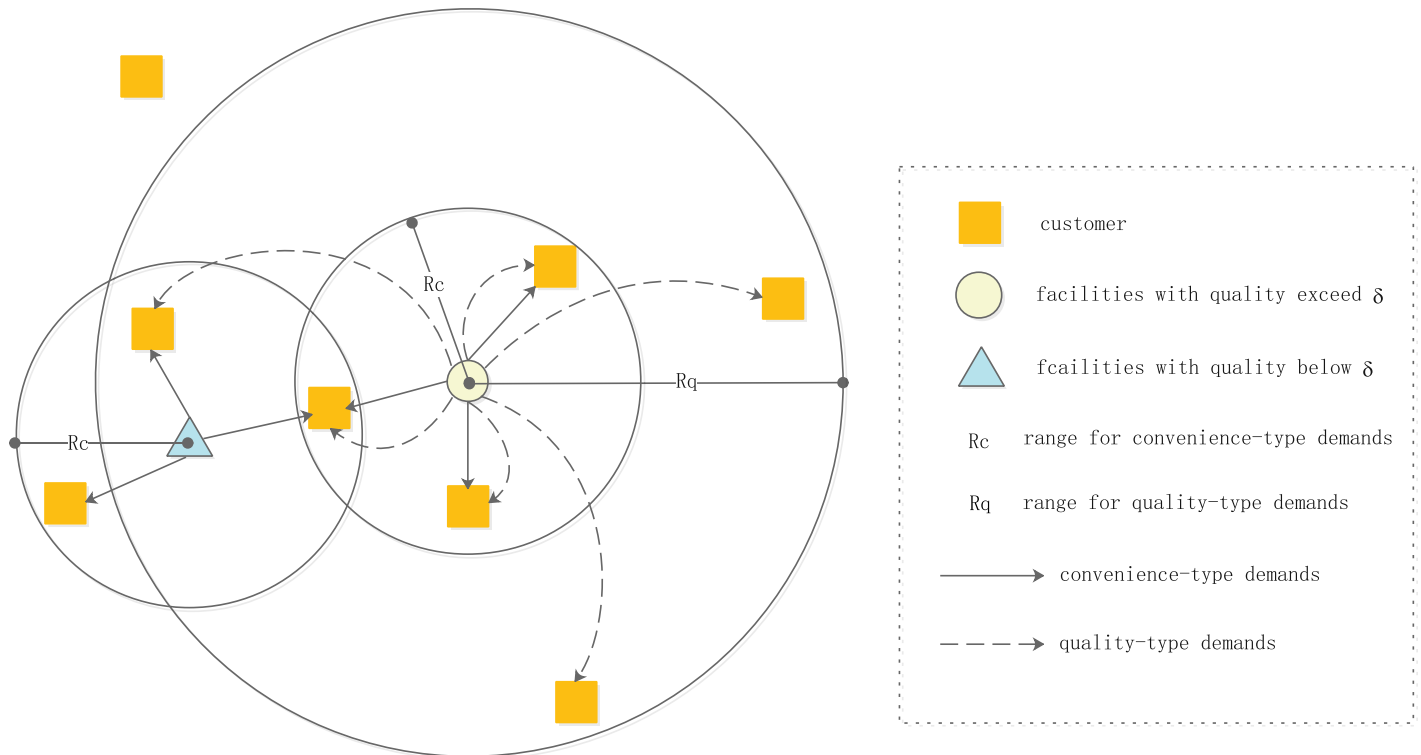


Fig 2. Patronage pattern from the perspective of facility.

<https://doi.org/10.1371/journal.pone.0273123.g002>

J : Union of set J_E and set J_N , i.e., $J = J_E \cup J_N$;

Q_j : Quality of facility at point $j, j \in J$;

δ : The quality threshold of facilities for quality-type demands;

R_c : The patronage range of customers for convenience-type demands;

R_q : The patronage range of customers for quality-type demands;

a_{ij}^C : The attractiveness of facility j to the convenience-type demands of customer i ;

a_{ij}^Q : The attractiveness of facility j to the quality-type demands of customer i ;

w_i^C : The convenience-type demands of customer $i \in I$;

w_i^Q : The quality-type demands of customer $i \in I$;

w_i : The total demands of customer i , that is, $w_i = w_i^C + w_i^Q, i \in I$;

W : The total demands of all customers, $W = \sum_{i \in I} w_i$;

γ_1 : The lower bound of the proportion of w_i^C in the total demands W , i.e. $0 \leq \gamma_1 \leq \sum_{i \in I} w_i^C / W$;

γ_2 : The upper bound of the proportion of w_i^C in the total demands W , i.e. $\sum_{i \in I} w_i^C / W \leq \gamma_2 \leq 1$;

f_j : The fixed cost of establishing a facility in the potential location $j \in J_N$;

d_{ij} : The distance between the customer i and the facility j ;

G: Total budget for the new entering company to build facilities;

$y_j, j \in J_N$: Binary variable, if the potential location j is selected to build a facility, then $y_j = 1$, otherwise $y_j = 0$.

From the definition of the convenience-type demands and the quality-type demands, we can define the attractiveness of facility j to the convenience-type (and quality-type) demands of customer i as follows:

$$a_{ij}^C = \begin{cases} \frac{Q_j}{1 + d_{ij}^2}, & \text{if } d_{ij} \leq R_c \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

$$a_{ij}^Q = \begin{cases} \frac{Q_j}{1 + d_{ij}^2}, & \text{if } d_{ij} \leq R_q \text{ and } Q_j \geq \delta \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

The goal of the new entrant is to maximize market share by setting new facilities in the competitive market under the given budget. Therefore, the competitive facility location problem of the new entering company can be established as the following two-layer programming model:

$$(OM) \quad \begin{aligned} Z_R &= \max P(y) \\ s.t. & \begin{cases} \sum_{j \in J_N} f_j y_j \leq G \\ y_j \in \{0, 1\}, \forall j \in J_N \end{cases} \end{aligned}$$

Here the outer-layer model (OM) is used to determine the location of new facilities with a limited budget in order to maximize the market share of the new entrant. According to the previous assumptions, we know that the proportion of $\sum_{i \in I} w_i^C / W$ is uncertain within the interval $[\gamma_1, \gamma_2]$. Therefore, the following inner-layer model (IM) can be used to deal with the uncertainty of convenience-type demands and quality-type demands of each customer:

$$(IM) \quad \begin{aligned} P(y) &= \min \sum_{i \in I} \left[w_i x_i \frac{\sum_{j \in J_N} a_{ij}^C y_j}{\sum_{j \in J_E} a_{ij}^C + \sum_{j \in J_N} a_{ij}^C y_j} + w_i (1 - x_i) \frac{\sum_{j \in J_N} a_{ij}^Q y_j}{\sum_{j \in J_E} a_{ij}^Q + \sum_{j \in J_N} a_{ij}^Q y_j} \right] \\ s.t. & \begin{cases} \gamma_1 \leq \sum_{i \in I} w_i x_i / W \leq \gamma_2 \\ 0 \leq x_i \leq 1, i \in I. \end{cases} \end{aligned}$$

Here x_i represents the ratio of convenience-type demands w_i^C to total demands w_i of customer i , i.e. $x_i = w_i^C / w_i$. And $\sum_{j \in J_N} a_{ij}^C y_j / \left(\sum_{j \in J_E} a_{ij}^C + \sum_{j \in J_N} a_{ij}^C y_j \right)$ represents the convenience-type demands captured by the new entrant from customer i .

4.2 Linearization of the original model

For brevity, let $h_i^C(y) := \frac{\sum_{j \in J_N} a_{ij}^C y_j}{\sum_{j \in J_E} a_{ij}^C + \sum_{j \in J_N} a_{ij}^C y_j}$, $h_i^Q(y) := \frac{\sum_{j \in J_N} a_{ij}^Q y_j}{\sum_{j \in J_E} a_{ij}^Q + \sum_{j \in J_N} a_{ij}^Q y_j}$, then (IM) is equivalent to:

$$P(y) = \sum_{i \in I} w_i h_i^Q(y) + \min \sum_{i \in I} w_i x_i [h_i^C(y) - h_i^Q(y)] \tag{3}$$

$$s.t. \begin{cases} \sum_{i \in I} w_i x_i \geq \gamma_1 W & (4a) \\ -\sum_{i \in I} w_i x_i \geq -\gamma_2 W & (4b) \\ -x_i \geq -1, i \in I & (4c) \\ x_i \geq 0, i \in I. & (4d) \end{cases}$$

Suppose the dual variables of constraints (4a–4c) are u_1, u_2 , and v_i , respectively. Then the dual problem of the minimization part of $P(y)$ is as follows:

$$\max (\gamma_1 W)u_1 - (\gamma_2 W)u_2 - \sum_{i \in I} v_i \tag{5}$$

$$s.t. \begin{cases} w_i(u_1 - u_2) - v_i \leq w_i(h_i^C(y) - h_i^Q(y)), \forall i \in I & (6a) \\ u_1, u_2 \geq 0, v_i \geq 0, \forall i \in I & (6b) \end{cases}$$

Therefore, the original two-layer model can be rewritten as a single-layer model as follows:

$$Z_R = \max \sum_{i \in I} w_i h_i^Q(y) + (\gamma_1 W)u_1 - (\gamma_2 W)u_2 - \sum_{i \in I} v_i \tag{7}$$

$$s.t. \begin{cases} \sum_{j \in J_N} f_j y_j \leq G & (8a) \\ w_i(u_1 - u_2) - v_i \leq w_i(h_i^C(y) - h_i^Q(y)), \forall i \in I & (8b) \\ u_1, u_2 \geq 0, v_i \geq 0, \forall i \in I & (8c) \\ y_j \in \{0, 1\}, \forall j \in J_N & (8d) \end{cases}$$

Since $h_i^C(y)$ and $h_i^Q(y)$ are nonlinear with respect to y_j , the above model is difficult to solve. Therefore, we linearize these two expressions by introducing two new decision variables α_i^C, α_i^Q , let $\alpha_i^C = h_i^C(y), \alpha_i^Q = h_i^Q(y)$.

$$\alpha_i^C = \frac{\sum_{j \in J_N} a_{ij}^C y_j}{\sum_{j \in J_E} a_{ij}^C + \sum_{j \in J_N} a_{ij}^C y_j}, \alpha_i^Q = \frac{\sum_{j \in J_N} a_{ij}^Q y_j}{\sum_{j \in J_E} a_{ij}^Q + \sum_{j \in J_N} a_{ij}^Q y_j} \tag{9}$$

From Eq (9), it can be seen that if $\sum_{j \in J_N} a_{ij}^C y_j = 0$ (or $\sum_{j \in J_N} a_{ij}^Q y_j = 0$), then $\alpha_i^C = 0$ (or $\alpha_i^Q = 0$), otherwise there are

$$\left(\sum_{j \in J_E} a_{ij}^C\right)\alpha_i^C + \left(\sum_{j \in J_N} a_{ij}^C y_j\right)\alpha_i^C = \sum_{j \in J_N} a_{ij}^C y_j \tag{10}$$

$$\left(\sum_{j \in J_E} a_{ij}^Q\right)\alpha_i^Q + \left(\sum_{j \in J_N} a_{ij}^Q y_j\right)\alpha_i^Q = \sum_{j \in J_N} a_{ij}^Q y_j \tag{11}$$

Now introduce two other decision variables $\beta_{ij}^C, \beta_{ij}^Q, i \in I, j \in J_N$, let

$$\beta_{ij}^C = \alpha_i^C y_j, \beta_{ij}^Q = \alpha_i^Q y_j \tag{12}$$

Since y_j is a binary variable, if $y_j = 0$ there are $\beta_{ij}^C = 0$ and $\beta_{ij}^Q = 0$, otherwise there are $\beta_{ij}^C = \alpha_i^C$ and $\beta_{ij}^Q = \alpha_i^Q$.

Therefore, the original model can be linearizing as follows:

$$Z_R = \max \sum_{i \in I} w_i \alpha_i^Q + (\gamma_1 W) u_1 - (\gamma_2 W) u_2 - \sum_{i \in I} v_i \tag{13}$$

$$(LM) \quad s.t. \quad \left\{ \begin{array}{l} \sum_{j \in J_N} f_j y_j \leq G \tag{14a} \\ w_i(u_1 - u_2) - v_i \leq w_i(\alpha_i^C - \alpha_i^Q), \forall i \in I \tag{14b} \\ \left(\sum_{j \in J_E} a_{ij}^C\right)\alpha_i^C + \sum_{j \in J_N} a_{ij}^C \beta_{ij}^C = \sum_{j \in J_N} a_{ij}^C y_j, \forall i \in I \tag{14c} \\ \left(\sum_{j \in J_E} a_{ij}^Q\right)\alpha_i^Q + \sum_{j \in J_N} a_{ij}^Q \beta_{ij}^Q = \sum_{j \in J_N} a_{ij}^Q y_j, \forall i \in I \tag{14d} \\ \alpha_i^C \leq M \sum_{j \in J_N} a_{ij}^C y_j, \forall i \in I \tag{14e} \\ \alpha_i^Q \leq M \sum_{j \in J_N} a_{ij}^Q y_j, \forall i \in I \tag{14f} \\ 0 \leq \alpha_i^C - \beta_{ij}^C \leq (1 - y_j), \forall i \in I, j \in J_N \tag{14g} \\ 0 \leq \alpha_i^Q - \beta_{ij}^Q \leq (1 - y_j), \forall i \in I, j \in J_N \tag{14h} \\ \beta_{ij}^C \leq y_j, \forall i \in I, j \in J_N \tag{14i} \\ \beta_{ij}^Q \leq y_j, \forall i \in I, j \in J_N \tag{14j} \\ u_1, u_2 \geq 0, v_i, \alpha_i^C, \alpha_i^Q, \beta_{ij}^C, \beta_{ij}^Q \geq 0, \forall i \in I, j \in J_N \tag{14k} \\ y_j \in \{0, 1\}, \forall j \in J_N \tag{14l} \end{array} \right.$$

Here M is a sufficiently large positive number. Constraint (14a) indicates that the total cost of all facilities does not exceed the budget. (14b) are the dual constraints on variables x_i in (4d). Constraints (14c–14f) guarantee that Eq (9) holds, and constraints (14g–14j) ensure that Eq (12) is true. Since this linearization model (LM) is a mixed binary linear programming, any commercial optimization software (such as Gurobi, Cplex, etc.) can solve medium-scale problems.

5 Ranking-based imbedding algorithm

For large-scale problems, current commercial optimization software cannot solve them well. Therefore, this section designs a ranking-based heuristic algorithm to solve large-scale problems.

Let $p_i := w_i[h_i^Q(y) - h_i^C(y)]$ and $\lambda_1 := \gamma_1 W, \lambda_2 := \gamma_2 W$, the minimization part of $P(y)$ is actually a generalized continuous knapsack problem (GCKP).

$$(GCKP) \quad \begin{aligned} & \max \sum_{i=1}^n p_i x_i \\ & s.t. \begin{cases} \lambda_1 \leq \sum_{i=1}^n w_i x_i \leq \lambda_2 \\ 0 \leq x_i \leq 1, i = 1, 2, \dots, n. \end{cases} \end{aligned}$$

where $w_i > 0, i = 1, 2, \dots, n$ and $0 \leq \lambda_1 < \lambda_2 \leq \sum_{i=1}^n w_i$. Please note that we call the problem (GCKP) ‘general’ not only it has upper and lower bounds, but also because the profit p_i may be negative. When the item $\sum_{i=1}^n w_i x_i$ has only an upper bound and all profits $p_i, i = 1, 2, \dots, n$ are non-negative, this problem is a continuous knapsack problem [30].

Reorder the index i so that the sorted index $k_i (i = 1, 2, \dots, n)$ satisfies:

$$\frac{p_{k_1}}{w_{k_1}} \geq \frac{p_{k_2}}{w_{k_2}} \geq \dots \geq \frac{p_{k_s}}{w_{k_s}} \geq 0 > \frac{p_{k_{s+1}}}{w_{k_{s+1}}} \geq \dots \geq \frac{p_{k_n}}{w_{k_n}}. \tag{15}$$

The optimal solution of (GCKP) can be determined by the following theorem.

Theorem. In addition to the key item s defined in (15), the other two key items s_1 and s_2 are defined as follows:

$$s_1 = \max \left\{ j : \sum_{i=1}^j w_{k_i} \leq \lambda_1 \right\}, s_2 = \min \left\{ j : \sum_{i=1}^j w_{k_i} > \lambda_2 \right\}. \tag{16}$$

Then the optimal solution x^* of (GCKP) is

(Case 1) If $s < s_1$, then

$$x_{k_i}^* = \begin{cases} 1, & i = 1, 2, \dots, s_1 - 1 \\ (\lambda_1 - \sum_{i=1}^{s_1-1} w_{k_i}) / w_{k_{s_1}}, & i = s_1 \\ 0, & i = s_1 + 1, \dots, n \end{cases} \tag{17}$$

(Case 2) If $s_1 \leq s < s_2$, then

$$x_{k_i}^* = \begin{cases} 1, & i = 1, 2, \dots, s \\ 0, & i = s + 1, \dots, n \end{cases} \tag{18}$$

(Case 3) If $s \geq s_2$, then

$$x_{k_i}^* = \begin{cases} 1, & i = 1, 2, \dots, s_2 - 1 \\ (\lambda_2 - \sum_{i=1}^{s_2-1} w_{k_i}) / w_{k_{s_2}}, & i = s_2 \\ 0, & i = s_2 + 1, \dots, n \end{cases} \tag{19}$$

Proof: (Case 1) If $s < s_1$, there is $0 > \frac{p_{k_{s_1}}}{w_{k_{s_1}}} \geq \dots \geq \frac{p_{k_n}}{w_{k_n}}$. Then for any feasible solution x that satisfies $\lambda_1 \leq \sum_{i=1}^n w_i x_i \leq \lambda_2$ and $0 \leq x_i \leq 1, i = 1, 2, \dots, n$. We have

$$\begin{aligned} \sum_{i=1}^n p_{k_i} x_{k_i} &= \sum_{i=1}^{s_1-1} p_{k_i} x_{k_i} + \sum_{i=s_1}^n \frac{p_{k_i}}{w_{k_i}} (w_{k_i} x_{k_i}) \leq \sum_{i=1}^{s_1-1} p_{k_i} x_{k_i} + \frac{p_{k_{s_1}}}{w_{k_{s_1}}} \sum_{i=s_1}^n w_{k_i} x_{k_i} \\ &\leq \sum_{i=1}^{s_1-1} p_{k_i} x_{k_i} + \frac{p_{k_{s_1}}}{w_{k_{s_1}}} (\lambda_1 - \sum_{i=1}^{s_1-1} w_{k_i} x_{k_i}) = \frac{p_{k_{s_1}}}{w_{k_{s_1}}} \lambda_1 + \sum_{i=1}^{s_1-1} (\frac{p_{k_i}}{w_{k_i}} - \frac{p_{k_{s_1}}}{w_{k_{s_1}}}) w_{k_i} x_{k_i} \\ &\leq \frac{p_{k_{s_1}}}{w_{k_{s_1}}} \lambda_1 + \sum_{i=1}^{s_1-1} (\frac{p_{k_i}}{w_{k_i}} - \frac{p_{k_{s_1}}}{w_{k_{s_1}}}) w_{k_i} = \sum_{i=1}^{s_1-1} p_{k_i} + \frac{p_{k_{s_1}}}{w_{k_{s_1}}} (\lambda_1 - \sum_{i=1}^{s_1-1} w_{k_i}) = \sum_{i=1}^n p_{k_i} x_{k_i}^* \end{aligned}$$

The first inequality is due to $\frac{p_{k_i}}{w_{k_i}} \leq \frac{p_{k_{s_1}}}{w_{k_{s_1}}}, \forall i = s_1, \dots, n$. The second inequality is due to $\frac{p_{k_{s_1}}}{w_{k_{s_1}}} < 0$ and $\lambda_1 - \sum_{i=1}^{s_1-1} w_{k_i} x_{k_i} \leq \sum_{i=s_1}^n w_{k_i} x_{k_i}$. The third inequality is due to $(\frac{p_{k_i}}{w_{k_i}} - \frac{p_{k_{s_1}}}{w_{k_{s_1}}}) w_{k_i} > 0, \forall i = 1, \dots, s_1 - 1$ and $x_{k_i} \leq 1$. Therefore, x^* in (17) is the optimal solution of (GCKP) when $s < s_1$.

(Case 2) If $s_1 \leq s < s_2$, the result is obvious.

(Case 3) It can be proved in a similar way to Case 1.

This theorem is proved by us and is not quoted from any other papers, it is a generalization of the optimal solution theorem for the continuous knapsack problem in [30]. The form of this optimal solution is not very intuitive, but the result obtained after combining w_{k_i} becomes very intuitive.

$$\sum_{i=1}^n w_{k_i} x_{k_i}^* = \begin{cases} \lambda_1, & \sum_{i=1}^s w_{k_i} < \lambda_1 \\ \sum_{i=1}^s w_{k_i}, & \lambda_1 \leq \sum_{i=1}^s w_{k_i} \leq \lambda_2 \\ \lambda_2, & \sum_{i=1}^s w_{k_i} > \lambda_2 \end{cases} \tag{20}$$

According to the result of Eq (20) and the non-increment of x_{k_i} , we propose the iterative algorithm for solving the (GCKP) as follows:

GCKP-A:

Step 1: Let $J_1 := \{j : p_j \geq 0\}, J_2 := \{j : p_j < 0\}, \bar{W} := \sum_{j \in J_1} w_j$. Set $x_j^* := 1$ for $j \in J_1$ and $x_j^* := 0$ for $j \in J_2$.

Step 2: While $\bar{W} < \lambda_1$:

Find the index j' such that $\frac{p_{j'}}{w_{j'}} = \max\left\{\frac{p_j}{w_j} : j \in J_2\right\}$.

If $w_{j'} \leq \lambda_1 - \bar{W}$, then set $x_{j'}^* := 1$, $\bar{W} := \bar{W} + w_{j'}$. Else set $x_{j'}^* := (\lambda_1 - \bar{W})/w_{j'}$, $\bar{W} := \lambda_1$.

Update $J_2 := J_2 \setminus \{j'\}$.

End while.

Step 3: While $\bar{W} > \lambda_2$:

Find the index j' such that $\frac{p_{j'}}{w_{j'}} = \min\left\{\frac{p_j}{w_j} : j \in J_1\right\}$.

If $w_{j'} \leq \bar{W} - \lambda_2$, then set $x_{j'}^* := 0$, $\bar{W} := \bar{W} - w_{j'}$. Else set $x_{j'}^* := 1 - (\bar{W} - \lambda_2)/w_{j'}$, $\bar{W} := \lambda_2$.

Update $J_1 := J_1 \setminus \{j'\}$.

End while.

Step 4: Output the optimal solution x^* and the objective value $\sum_{j=1}^n p_j x_j^*$.

We know that the inner-layer model (IM) can be accurately solved by (GCKP-A) for any given location scheme of the new entering firm, so the remaining problem is to efficiently solve the outer-layer model (OM). Fernández et al. [16] proposed a new rule: Pareto-Huff customer choice rule, and used the ranking-based heuristic strategy (RDOA) to solve the corresponding competitive facility location problem. By setting a different rank rule, we [31] proposed an improved ranking-based algorithm for the competitive facility location problem. Compared with other heuristic algorithms (such as genetic algorithm, simulated annealing algorithm, ant colony algorithm, etc.), the ranking-based heuristic algorithm performs very well on the facility location problem. Therefore, by embedding (GCKP-A) and the 2-opt strategy, we proposed a ranking-based embedding algorithm (REA) to solve the model of this paper. The flowchart of the algorithm (REA) is shown in Fig 3.

Ranking-based Embedding Algorithm (REA)

Step 1. (Initialization of the algorithm):

Input parameters of the model, set the iterate number L and set initial ranking scores $r_j :=$

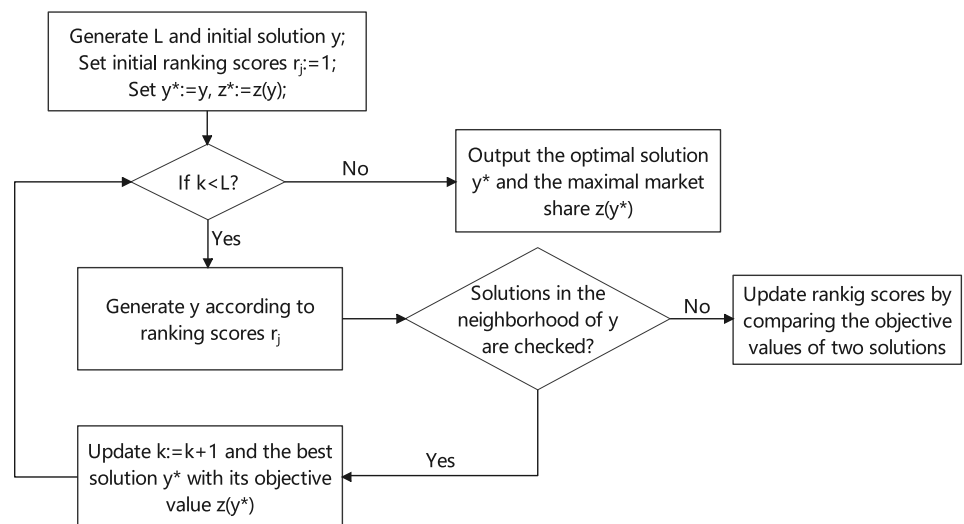


Fig 3. Flowchart of the algorithm (REA).

<https://doi.org/10.1371/journal.pone.0273123.g003>

1, $\forall j \in J_N$, set $\lambda_C := (\lambda_1 + \lambda_2)/2$. Randomly choose y^* such that $\sum_{j \in J_N} f_j y_j^* \leq G$, calculate $P(y^*)$ by (GCKP-A). Let $k := 0$ and set $Z_R^* := P(y^*)$.

Step 2. (Ranking-based algorithm embedding (GCKP-A)):

While $k < L$:

Step 2.1 (Generate a solution y according to the ranking-based rule)

Let $y_j := 0, \forall j \in J_N$, calculate the probability p_j of each candidate location $j \in J_N$:

$$p_j := \frac{A_j r_j}{\sum_{j \in J_N} A_j r_j}, \forall j \in J_N, \text{ where } A_j = \lambda_C \sum_{i \in I} a_{ij}^C + (1 - \lambda_C) \sum_{i \in I} a_{ij}^Q$$

Repeat the roulette-wheel selection process to set $y_j := 1$ until $\sum_{j \in J_N} f_j y_j > G$.

Step 2.2 (Use 2-opt strategy to update the ranking scores)

Randomly choose $j' \in Y^1 := \{j \in J_N | y_j = 1\}$. Let

$$F' := \left\{ k \in (J_N \setminus Y^1) | f_k \leq G - \sum_{j \in J_N} f_j y_j + f_{j'} \right\}.$$

For each $\bar{j} \in F'$:

Let $\tilde{y} := y$. Set $\tilde{y}_{j'} := 0, \tilde{y}_{\bar{j}} := 1$, calculate $P(y)$ and $P(\tilde{y})$ by (GCKP-A).

If $P(y) > P(\tilde{y})$:

Set $r_j := r_j + y_j \cdot 2, r_{j'} := r_{j'} - \tilde{y}_{j'}$; Update $y^* := y, Z_R^* := P(y)$ when $P(y) > Z_R^*$;

Else:

Set $r_j := r_j + \tilde{y}_{j'} \cdot 2, r_{\bar{j}} := r_{\bar{j}} - y_{\bar{j}}$; Update $y^* := \tilde{y}, Z_R^* := P(\tilde{y})$ when $P(\tilde{y}) > Z_R^*$.

Update $r_j := r_j + 1, \forall j \in J_N$ if any $r_j = 0$.

End for.

$k := k + 1$.

End while

Step 3. Algorithm termination. Output the optimal solution y^* and the maximal market share Z_R^* .

In (REA), we embed (GCKP-A) to solve the inner-layer problem and use the 2-opt strategy to update the ranking score. Because (GCKP-A) is much faster than using other linear programming algorithms to solve the inner-layer problem, it can improve the computational efficiency of (REA). The 2-opt strategy can make the ranking results more comparable, so as to obtain a better combination of facility locations.

6 Computational experiments

6.1 Results for benchmark problems

By comparing the performance of our algorithm (REA) with the algorithm (RDOA-D) [16] on 40 benchmark problems (pmed1~pmed40) [32], this subsection demonstrates the effectiveness of (REA). Note that computational experiments in [16] prove that the algorithm (RDOA-D) is superior to the algorithm (RDOA) and the genetic algorithm (GA) for all parameters of the tested problem, so we only compare the performance of algorithm (REA) and algorithm (RDOA-D). The number of nodes in these benchmark instances ranges from 100 to 900. The distance d_{ij} between the pair of points i and j is the Euclidean distance. All experiments are performed in a MATLAB environment on a laptop computer with Intel(R) Core (TM) i7-1065G7 @1.30 GHz CPU, and 16.00 GB of RAM.

Table 1. Comparison of algorithm efficiency for solving (GCKP) (time unit: Second).

	<i>n</i> = 100	<i>n</i> = 200	<i>n</i> = 300	<i>n</i> = 400	<i>n</i> = 500	<i>n</i> = 600	<i>n</i> = 700	<i>n</i> = 800	<i>n</i> = 900	<i>n</i> = 1000	Mean
A1	0.0003	0.0003	0.0004	0.0002	0.0002	0.0002	0.0002	0.0004	0.0005	0.0005	0.00032
A2	0.1762	0.1899	0.1894	0.1859	0.1826	0.1925	0.1809	0.1857	0.2979	0.2896	0.2071

<https://doi.org/10.1371/journal.pone.0273123.t001>

First, for generalized continuous knapsack problems of different scales, we compare the efficiency of the algorithm (GCKP-A, A1 in Table 1) with the algorithm used for linear programming algorithm in Gurobi 9.1.1 (Linear-A, A2 in Table 1). For each example, the parameters p_j and w_j are randomly generated from intervals $[-100, 100]$ and $[10, 100]$, respectively. The optimal solutions obtained by the two algorithms are the same, and the CPU times used are shown in Table 1. It can be seen from Table 1 that the average CPU time used by (Linear-A) is more than 640 times that of (GCKP-A), so the calculation efficiency of (GCKP-A) is significantly higher than (Linear-A) of Gurobi. Therefore, using (GCKP-A) in (REA) instead of linear algorithm of Gurobi can greatly improve the efficiency of the algorithm (REA).

Next, we test the optimality of the algorithm (REA) and compare the performance of Simulated Annealing Algorithm (SA). For the first 5 benchmark examples (pmed1~pmed5), we compare the results obtained by (REA) with the exact results (Market Share (MS)) obtained by solving the linearized model (LM). In each example, suppose that the first 10 locations are occupied by existing facilities, and the next $|J_N|$ locations are candidate facility locations for the new entrant. The quality Q_j of facility j is randomly generated as an integer number in $[10, 100]$. The customer demands of node i set to be $w_i = 100/i$ and the fixed cost f_j of opening a new facility at candidate node j randomly generated as an integer number in $[200, 500]$. Other parameters are as follows: $R_c = 60, R_q = 100, \delta = 80, \gamma_1 = 0.4, \gamma_2 = 0.6, G = 2000$. For each example, (REA) and (SA) were run 20 times, and the results are listed in Table 2. The Gap is defined as follows: $Gap = \frac{MS - Mean}{MS} \times 100\%$. From Table 2, we can find that the maximum gap between the average market share obtained from the algorithm (REA) and the exact result is no more than 0.35%. This verifies the optimality of the proposed algorithm (REA). On the other hand,

Table 2. Optimality of the algorithm (REA).

No.	$ J_N $	Exact	(REA)				(SA)			
		MS	Max.	Min.	Mean	Gap	Max.	Min.	Mean	Gap
#1	30	86.0944	86.0944	86.0944	86.0944	0%	86.0944	85.0634	85.8367	0.3%
	40	100.1909	100.1909	100.1909	100.1909	0%	100.1909	97.3497	99.8464	0.34%
	50	101.5025	101.5025	101.5025	101.5025	0%	101.5025	96.1879	99.8296	1.65%
#2	30	102.9141	102.9141	102.9141	102.9141	0%	102.9141	102.8386	102.9028	0.01%
	40	127.7107	127.7107	127.7107	127.7107	0%	127.7107	124.7100	126.6604	0.82%
	50	129.9127	129.9126	127.2366	129.4630	0.35%	129.9126	124.7341	129.6025	0.24%
#3	30	88.6559	88.6559	88.6559	88.6559	0%	88.6559	85.9735	88.3029	0.40%
	40	111.1247	111.1242	110.1094	111.0734	0.05%	111.1242	106.5701	110.8965	0.21%
	50	101.6469	101.6469	100.2887	101.5790	0.07%	101.6469	94.1251	99.8916	1.73%
#4	30	118.1794	118.1794	118.0909	118.1219	0.05%	118.1794	118.0105	118.0612	0.10%
	40	122.4177	122.4177	122.4177	122.4177	0%	122.4177	120.7804	122.2540	0.13%
	50	70.3602	70.3602	70.3602	70.3602	0%	70.3602	63.7387	69.7246	0.90%
#5	30	80.3325	80.3325	80.3325	80.3325	0%	80.3325	79.5055	80.2085	0.15%
	40	104.0804	104.0804	104.0804	104.0804	0%	104.0804	99.9589	103.8743	0.20%
	50	82.1296	82.1296	80.8194	82.0641	0.08%	82.1296	77.8385	81.7646	0.44%

<https://doi.org/10.1371/journal.pone.0273123.t002>

Table 3. Performance comparison of (REA) and (RDOA-D).

(#:G, I , J _N)	(REA)		(RDOA-D)		(#:G, I , J _N)	(REA)		(RDOA-D)	
	Mean	Std.	Mean	Std.		Mean	Std.	Mean	Std.
1:2000,100,60	93.966	0.185	93.391	0.676	21:3000,500,60	231.714	0.231	229.980	1.769
2:2000,100,60	118.544	0.119	118.513	0.077	22:3000,500,60	303.405	0.131	303.088	0.416
3:2000,100,60	104.674	1.304	104.845	1.180	23:3000,500,60	330.032	0.000	329.904	0.299
4:2000,100,60	104.694	0.178	103.128	2.418	24:3000,500,60	275.718	0.000	275.592	0.163
5:2000,100,60	144.267	2.530	141.658	2.801	25:3000,500,60	334.615	0.537	333.547	0.816
6:2000,200,80	151.319	0.520	150.159	0.737	26:3000,600,80	276.799	0.000	276.541	0.557
7:2000,200,80	179.535	1.470	180.449	0.783	27:3000,600,80	332.570	0.000	331.917	0.630
8:2000,200,80	234.200	0.000	234.200	0.000	28:3000,600,80	318.632	0.000	318.462	0.225
9:2000,200,80	153.915	0.906	150.949	1.941	29:3000,600,80	301.130	0.360	300.459	0.443
10:2000,200,80	225.614	0.000	224.554	1.367	30:3000,600,80	315.423	0.798	314.436	1.294
11:2500,300,80	239.610	0.529	237.713	1.160	31:3500,700,80	295.503	0.000	294.107	1.279
12:2500,300,80	247.627	0.000	247.512	0.182	32:3500,700,80	305.194	0.161	303.251	0.912
13:2500,300,80	206.564	0.105	206.491	0.172	33:3500,700,80	320.562	0.000	319.979	0.506
14:2500,300,80	252.478	0.107	252.074	0.719	34:3500,700,80	383.017	0.107	382.329	0.500
15:2500,300,80	232.803	0.638	232.243	1.182	35:3500,800,80	348.047	0.000	347.483	0.499
16:2500,400,100	288.557	0.134	288.621	0.000	36:3500,800,100	339.951	0.584	336.575	1.323
17:2500,400,100	230.678	1.144	229.300	1.023	37:3500,800,100	349.700	0.279	347.553	1.144
18:2500,400,100	302.579	0.570	301.718	0.579	38:3500,900,100	337.952	0.874	335.746	0.788
19:2500,400,100	314.042	0.000	311.155	2.262	39:3500,900,100	334.147	0.278	332.247	0.892
20:2500,400,100	306.630	1.137	302.049	1.548	40:3500,900,100	389.986	0.681	387.673	0.420

<https://doi.org/10.1371/journal.pone.0273123.t003>

the Gap of the Simulated Annealing Algorithm is larger than the Gap of (REA) for almost every example. Therefore, we can conclude that the performance of (REA) is better than (SA).

Finally, we compared the performance of the algorithm (REA) with the algorithm (RDOA-D) using 40 benchmark instances. Similarly, assuming that in each example, the first 10 locations are occupied by existing facilities, the next $|J_N|$ locations are candidate facility locations for the new entrant. Except for the parameters given in Table 3, all other parameters are the same as those used in Table 2. We performed 20 calculations for each instance, and the mean and standard deviation of the market shares obtained are shown in Table 3. The larger mean and smaller standard deviation of the two algorithms are shown in bold.

It can be clearly seen from Table 3 that the algorithm (REA) is superior to the algorithm RDOA-D in both the mean and standard deviation of the market share.

6.2 A quasi-real example

In order to investigate the effects of different parameters in the presented competitive facility location model. This subsection discusses a quasi-real example based on the 49-nodes data set described in Daskin [33]. This data set consists of the capitals of the continental United States and Washington, DC.

The customer demands d_i is proportional to the population of the state (i) in 1990. The opening cost f_j of the potential facility location $j \in J_N$ also comes from Daskin's data set. In this example, we generated a set of random numbers within [10, 120] to represent the facility quality Q_j at each point $j \in J$. Now, there are 5 facilities located in the states of California, Kentucky, New York, Texas, and Wyoming, respectively. The qualities of these 5 facilities are 106, 98, 107, 45, and 106 respectively. To avoid being too close, we multiply the Euclidean distance

Table 4. Location scheme of the new entrant ($\gamma_1 = 0.4, \gamma_2 = 0.6$).

Location	Demand type	Customers Served
Georgia	Convenience	Alabama, Florida, Georgia, South Carolina, Tennessee
Indiana	Quality/ Convenience	Alabama, Arkansas, Georgia, Illinois, Indiana, Iowa, Kentucky, Michigan, Missouri, Ohio, South Carolina, Tennessee, West Virginia, Wisconsin
Missouri	Convenience	Arkansas, Illinois, Iowa, Kansas, Missouri

<https://doi.org/10.1371/journal.pone.0273123.t004>

between any two points by 10, and then round it to d_{ij} . Other parameters are defined as follows: $G = 200000$, $\delta = 100$, $R_c = 40$, $R_q = 80$, $\gamma_1 = 0.4$, $\gamma_2 = 0.6$.

The best solution to the problem of the new entrant's competitive facility location is to open three facilities in Georgia, Indiana, and Missouri. The newly-entered company's market share from its three facilities is 7821.1. The locations of the new facilities and customers served are shown in Table 4.

Let γ_1 change from 0.1 to 0.9 with a step size of 0.1 to test the impact of γ_1, γ_2 on the location scheme while maintaining $\gamma_2 > \gamma_1$. We found that when $\gamma_1 \leq 0.8$, no matter what value γ_2 takes in $[\gamma_1, 1]$, the location scheme is exactly the same as that shown in Table 4. Only when $\gamma_1 \geq 0.9$, the location scheme will be different. In this case, the new company will build three facilities in Georgia, Indiana, and Pennsylvania. Table 5 shows the location scheme when $\gamma_1 = 0.9, \gamma_2 = 1$. The market share captured by the new entering firm is 8159.2, which is larger than the previous market share 7821.1. This reflects that the reduced degree of uncertainty can bring more market share to the new entrant.

In addition to γ_1, γ_2 , there are other important parameters in the model such as G, R_q, R_c, δ , etc. Now we analyze the influence of these parameters on the market share of the newly entered company.

Firstly, we analyze the impact of R_q on market share. Set the budget G to 200000, 250000, and 300000 respectively. For each value of G , let R_q change from 60 to 130 with a step of 10, we can get a curve representing the change of market share with R_q . The three curves with different budgets are shown in Fig 4.

It can be clearly seen from Fig 4 that for different budgets, with the increase of R_q , the market share first increases and then decreases. Due to the increase in R_q , the number of demand points that can be met by high-quality facilities increases, so the market share first increases with the increase in R_q . With the further increase of R_q , the coverage of existing high-quality facilities has also increased, leading to intensified competition among facilities, so the available market share of the newly entered firm has begun to decline.

Secondly, we analyze the impact of R_c on market share. We still set the budget G to 200000, 250000, and 300000 respectively. Then change R_c from 10 to 70 in step size of 10. The relationship curves between market share and R_c are shown in Fig 5.

It can be seen from Fig 5 that the market share increases with the increase of R_c . When R_c changes from 10 to 20, the demand points covered by low-quality facilities are unchanged, so

Table 5. Location scheme of the new entrant ($\gamma_1 = 0.9, \gamma_2 = 1$).

Location	Demand type	Customers Served
Georgia	Convenience	Alabama, Florida, Georgia, South Carolina, Tennessee
Indiana	Quality/ Convenience	Alabama, Arkansas, Georgia, Illinois, Indiana, Iowa, Kentucky, Michigan, Missouri, Ohio, South Carolina, Tennessee, West Virginia, Wisconsin
Pennsylvania	Convenience	Delaware, Maryland, New Jersey, New York, Pennsylvania, Virginia

<https://doi.org/10.1371/journal.pone.0273123.t005>

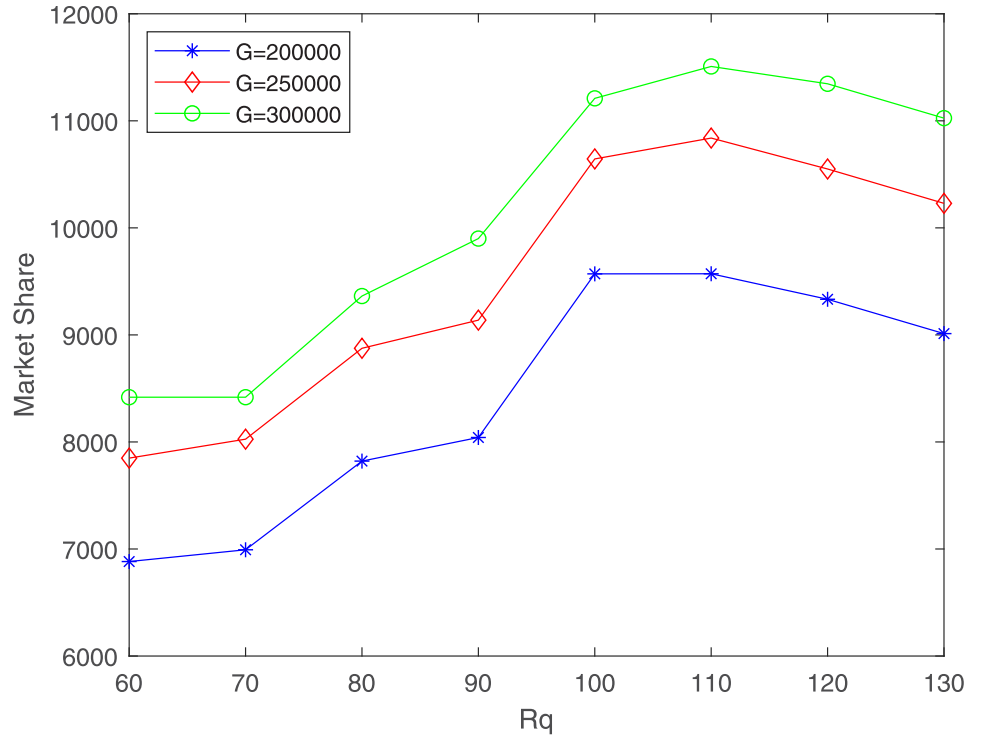


Fig 4. The relationship between market share and R_q ($R_c = 40, \delta = 100, \gamma_1 = 0.4, \gamma_2 = 0.6$).

<https://doi.org/10.1371/journal.pone.0273123.g004>

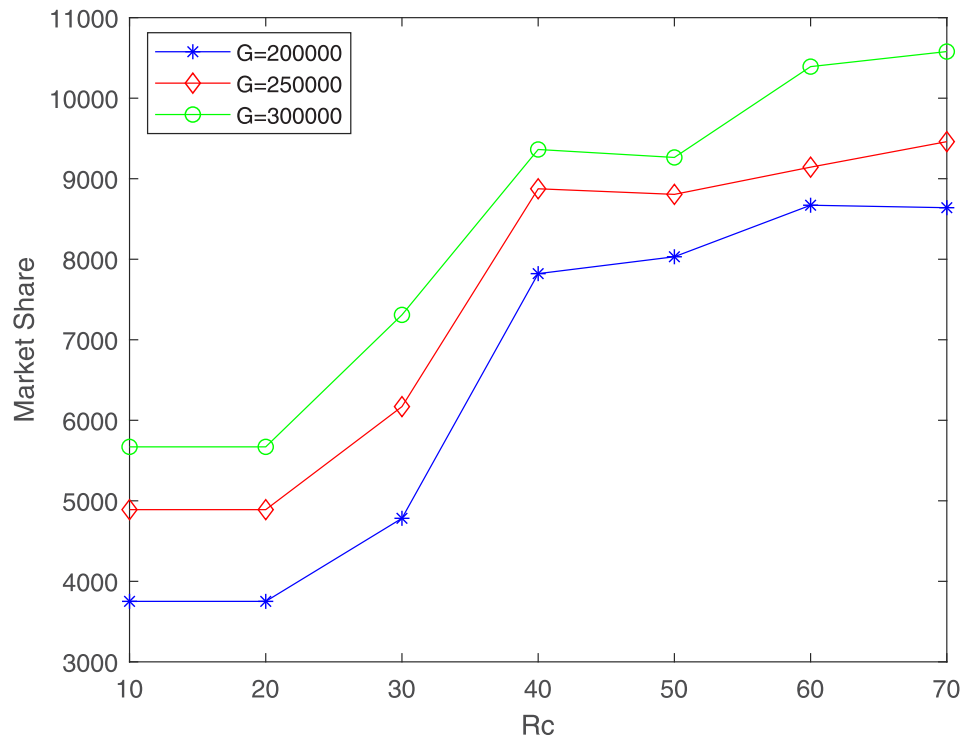


Fig 5. The relationship between market share and R_c ($R_q = 80, \delta = 100, \gamma_1 = 0.4, \gamma_2 = 0.6$).

<https://doi.org/10.1371/journal.pone.0273123.g005>

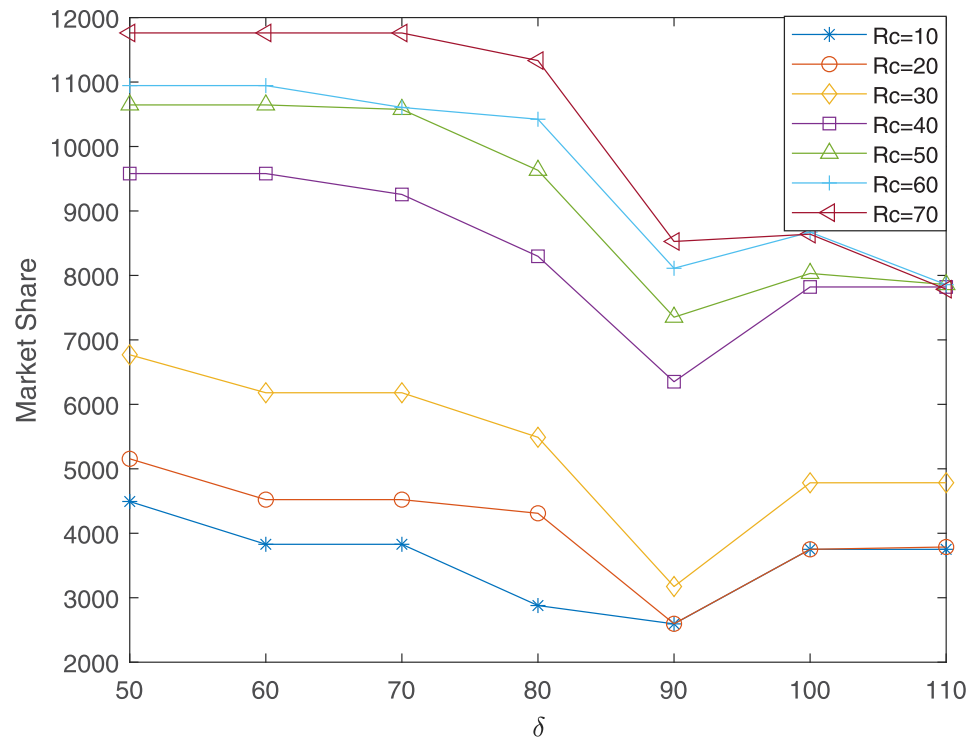


Fig 6. The relationship between market share and δ ($R_q = 80, \gamma_1 = 0.4, \gamma_2 = 0.6$).

<https://doi.org/10.1371/journal.pone.0273123.g006>

the market share remains the same. When R_c changes from 20 to 40, the demand points covered by low-quality facilities increase, so the market share increases significantly. When R_c further increases, competition among low-quality facilities begins to emerge, and therefore, the growth of market share slows down.

Thirdly, we analyze the impact of δ on market share. Let $R_q = 80$ and δ changes from 50 to 110 in step size of 10, for 7 different values of R_c , the relationship curves between market share and δ are shown in Fig 6.

It can be seen from Fig 6 that when δ changes from 50 to 90, the market share decreases as δ increases. This is because with the increase of δ , the number of facilities that meet the quality requirements higher than δ gradually decreases. When δ changes from 90 to 100, the market share increases as δ increases. This is due to the reduction in the number of existing facilities that meet the quality higher than δ , resulting in less competition among high-quality facilities. When δ changes from 100 to 110, the market share does not increase with the increase of δ . This is because when the δ further increases, the number of facilities that meet the quality requirements decreases significantly, so the market share that the new entrant can obtain remains unchanged or slightly decreases. For any value of R_c , when $\delta = 90$, the market share reaches the minimum value. In order to study the reason in-depth, we take $R_c = 40$ as an example, and the locations of $\delta = 80, 90, 100$ are listed in Table 6.

It can be seen from Table 6 that when δ increases, the number of high-quality facilities (Q_i underlined in Table 4) in the selected facilities gradually decreases. Denote the solution when $\delta = 80$ as $y_{\delta=80}^*$, then the market share corresponding to this solution when $\delta = 90$ is 4887. Similarly, denote the solution when $\delta = 100$ as $y_{\delta=100}^*$, then the market share corresponding to this solution when $\delta = 90$ is 5496.8.

Table 6. The location scheme for different δ ($R_c = 40, R_q = 80, \gamma_1 = 0.4, \gamma_2 = 0.6$).

$\delta = 80$			$\delta = 90$			$\delta = 100$		
Locations	Q_j	MS	Locations	Q_j	MS	Locations	Q_j	MS
Arkansas	91	8297.5	Arkansas	91	6351.1	Georgia	36	7821.1
Indiana	111		Indiana	111		Indiana	111	
Virginia	82		Kansas	67		Missouri	24	

<https://doi.org/10.1371/journal.pone.0273123.t006>

Similarly, let $R_c = 40$ and δ changes from 50 to 110 in step size of 10, for different values of R_q , the relationship curves between market share and δ are shown in Fig 7.

It can be seen in Fig 7 that when $\delta \leq 90$, the impact of R_q on the market share is not obvious. But when $\delta \geq 100$, there is a clear positive correlation between market share and R_q . This is because as δ increases, the number of existing facilities meeting high-quality requirements decreases. Therefore, for the new entrant, the market share brought about by choosing high-quality facilities will increase with the increase of R_q .

Finally, we analyzed the impact of the budget G on the location of the new entrant’s facilities, and the results are listed in Table 7. The last row in Table 7 indicates the market share of the new entrant.

It can be seen from Table 7 that the market share of the new entrant increases as the budget increases. An interesting phenomenon is that for a multi-stage budget situation, it is also possible to obtain the overall optimal solution. For example, suppose that the first stage budget is 150000, and the second stage budget is 100000. We can choose Georgia and Indiana in the first stage, and then choose Iowa and South Dakota in the second stage. This is exactly the optimal location scheme under the total budget of 250000.

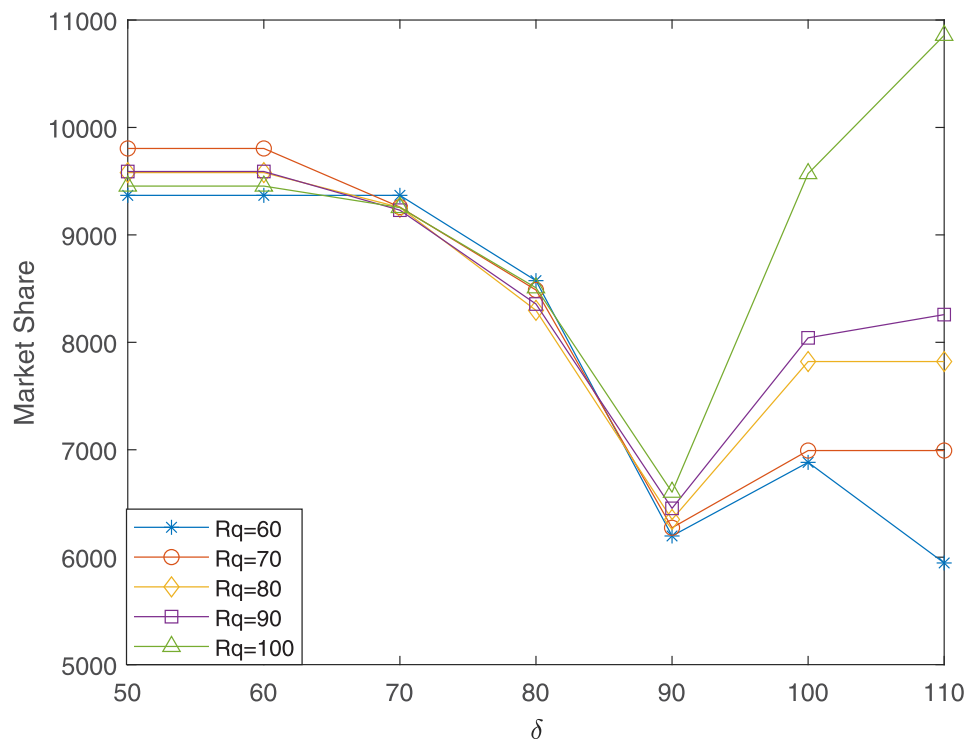


Fig 7. The relationship between market share and δ ($R_c = 40, \gamma_1 = 0.4, \gamma_2 = 0.6$).

<https://doi.org/10.1371/journal.pone.0273123.g007>

Table 7. The location scheme for different G ($R_c = 40, R_q = 80, \delta = 100, \gamma_1 = 0.4, \gamma_2 = 0.6$).

G = 100000		G = 150000		G = 200000		G = 250000		G = 300000	
Locations	Q_j	Locations	Q_j	Locations	Q_j	Locations	Q_j	Locations	Q_j
Indiana	111	Georgia	36	Georgia	36	Georgia	36	Alabama	31
		Indiana	111	Indiana	111	Indiana	111	Indiana	111
				Missouri	24	Iowa	16	Iowa	16
						South Dakota	114	Ohio	51
								South Dakota	114
								West Virginia	52
4321.0		6224.7		7821.1		8875.0		9363.0	

<https://doi.org/10.1371/journal.pone.0273123.t007>

7 Conclusion

Observing that many customers often have both convenience and quality demands, we propose a new customer choice rule to describe this behavior. In fact, if the quality threshold δ is set to infinity, the customer choice rule based on convenience range proposed by Qi et al. [23] can be regarded as a special case of the rule in this paper. By considering the uncertainty of the proportion of convenience-type demands in the total demands, we proposed a two-layer robust model to study the competitive facility location problem.

The initial two-layer robust model is linearized as a mixed binary linear programming problem. This allows us to solve exactly medium size problems. For large size problems, we first prove a theorem for the optimal solution of the generalized continuous Knapsack problem. Then we propose an exact algorithm (GCKP-A) to solve the inner-layer model. By imbedding (GCKP-A) and 2-opt strategy into the framework of the improved ranking-based algorithm, we propose the heuristic algorithm (REA) for large size problems. On the one hand, the solution of the proposed algorithm (REA) is compared with the exact solution for different benchmark examples, which verifies the optimality of the proposed algorithm. On the other hand, the comparison of the performance of the algorithm (REA) and the algorithm (RDOA-D) reflects the superiority of the algorithm (REA).

A quasi-real example is used to illustrate the impact of different parameters on the market share of the new entrant. Four main conclusions are obtained through sensitivity analysis: (1) Only when the proportion of convenience-type demands is high, will it affect the robust location scheme; (2) The market share of the new entrant increases first and then decreases with the increase of R_q ; (3) The market share of the new entrant increases with the increase of R_c ; (4) The market share of the new entrant decreases first and then does not decrease with the increase of δ .

In this article, we only studied the competitive facility location problem in a static environment. Based on the classification of these two types of demands, we should consider the reaction of the competitors in the location of competitive facilities. The leader-follower competitive facility location problem based on the newly proposed customer choice rule deserves to be studied in future work.

Supporting information

S1 File. The data set of the 49-nodes example in subsection 6.2 is provided as S1 File. The first two columns of the data are latitudes and longitudes of the capitals of the continental United States plus Washington, DC. The third column denotes the existing facilities. The fourth column is the demand d_j . The fifth column denotes the quality of facilities. The sixth column represents the cost of opening a facility at position j . The last column is the name of the states. (XLS)

Author Contributions

Conceptualization: Wuyang Yu.

References

1. Santos-Peñate DR, Suárez-Vega R, Dorta-González P. The leader-follower location model. *Networks and Spatial Economics*. 2007; 7:45–61. <https://doi.org/10.1007/s11067-006-9007-2>
2. ReVelle CS, Eiselt HA, Daskin MS. A bibliography for some fundamental problem categories in discrete location science. *European Journal of Operational Research*. 2008; 184:817–848. <https://doi.org/10.1016/j.ejor.2006.12.044>
3. Plastria F. Static competitive facility location: An overview of optimisation approaches. *European Journal of Operational Research*. 2001; 129:461–470. [https://doi.org/10.1016/S0377-2217\(00\)00169-7](https://doi.org/10.1016/S0377-2217(00)00169-7)
4. Serra D, Eiselt HA, Laporte G, ReVelle CS. Market capture models under various customer-choice rules. *Environment and Planning B: Planning and Design*. 1999; 26:741–750. <https://doi.org/10.1068/b260741>
5. Hotelling H. Stability in competition. *The Economic Journal*. 1929; 39:41–57. <https://doi.org/10.2307/2224214>
6. Huff DL. Defining and estimating a trading area. *Journal of Marketing*. 1964; 28:34–38. <https://doi.org/10.1177/002224296402800307>
7. Farahani RZ, Rezapour S, Drezner T, Fallah S. Competitive supply chain network design: An overview of classifications, models, solution techniques and applications. *Omega*. 2014; 45:92–118. <https://doi.org/10.1016/j.omega.2013.08.006>
8. Drezner T, Drezner Z, Zerom D. Competitive facility location with random attractiveness. *Operations Research Letters*. 2018; 46:312–317. <https://doi.org/10.1016/j.orl.2018.02.008>
9. Fernández J, Tóth BG, Redondo JL, Ortigosa PM. The probabilistic customer's choice rule with a threshold attraction value: Effect on the location of competitive facilities in the plane. *Computers & Operations Research*. 2019; 101:234–249. <https://doi.org/10.1016/j.cor.2018.08.001>
10. Marianov V, Eiselt HA, Lüer-Villagra A. The follower competitive location problem with comparison-shopping. *Networks and Spatial Economics*. 2020; 20:367–393. <https://doi.org/10.1007/s11067-019-09481-6>
11. Ljubić I, Moreno E. Outer approximation and submodular cuts for maximum capture facility location problems with random utilities. *European Journal of Operational Research*. 2018; 266:46–56. <https://doi.org/10.1016/j.ejor.2017.09.023>
12. Ahmadi Z, Ghezavati V. Developing a new model for a competitive facility location problem considering sustainability using Markov chains. *Journal of Cleaner Production*. 2020; 273:122971. <https://doi.org/10.1016/j.jclepro.2020.122971>
13. Mahmoodjanloo M, Tavakkoli-Moghaddam R, Baboli A, Jamiri A. A multi-modal competitive hub location pricing problem with customer loyalty and elastic demand. *Computers & Operations Research*. 2020; 123:105048. <https://doi.org/10.1016/j.cor.2020.105048>
14. Mai T, Lodi A. A multicut outer-approximation approach for competitive facility location under random utilities. *European Journal of Operational Research*. 2020; 284:874–881. <https://doi.org/10.1016/j.ejor.2020.01.020>
15. Aboolian R, Berman O, Krass D. Optimizing facility location and design. *European Journal of Operational Research*. 2021; 289:31–43. <https://doi.org/10.1016/j.ejor.2020.06.044>
16. Fernández P, Pelegrín B, Lančinskas A, Žilinskas J. Exact and heuristic solutions of a discrete competitive location model with Pareto-Huff customer choice rule. *Journal of Computational and Applied Mathematics*. 2021; 385:113200. <https://doi.org/10.1016/j.cam.2020.113200>
17. Fahimi K, Seyedhosseini SM, Makui A. Simultaneous competitive supply chain network design with continuous attractiveness variables. *Computers & Industrial Engineering*. 2017; 107:235–250. <https://doi.org/10.1016/j.cie.2017.03.020>
18. Zhou W, Zheng Y, Huang W. Competitive advantage of qualified WEEE recyclers through EPR legislation. *European Journal of Operational Research*. 2017; 257:641–655. <https://doi.org/10.1016/j.ejor.2016.07.050>
19. Cai K, He Z, Lou Y, He S. Risk-aversion information in a supply chain with price and warranty competition. *Annals of Operations Research*. 2020; 287:61–107. <https://doi.org/10.1007/s10479-019-03386-w>
20. Shakouhi F, Tavakkoli-Moghaddam R, Baboli A, Bozorgi-Amiri A. A competitive pharmaceutical supply chain under the marketing mix strategies and product life cycle with a fuzzy stochastic demand. *Annals of Operations Research*. 2021. <https://doi.org/10.1007/s10479-021-04073-5>

21. Sazvar Z, Zokaei M, Tavakkoli-Moghaddam R, Salari SA, Nayeri S. Designing a sustainable closed-loop pharmaceutical supply chain in a competitive market considering demand uncertainty, manufacturer's brand and waste management. *Annals of Operations Research*. 2021. <https://doi.org/10.1007/s10479-021-03961-0>
22. Yu Z, Khan SAR. Green supply chain network optimization under random and fuzzy environment. *International Journal of Fuzzy Systems*. 2022; 24:1170–1181. <https://doi.org/10.1007/s40815-020-00979-7>
23. Qi M, Xia M, Zhang Y, Miao L. Competitive facility location problem with foresight considering service distance limitations. *Computers & Industrial Engineering*. 2017; 112:483–491. <https://doi.org/10.1016/j.cie.2017.04.024>
24. Yu W. A leader-follower model for discrete competitive facility location problem under the partially proportional rule with a threshold. *PLOS ONE*. 2019; 14:e0225693. <https://doi.org/10.1371/journal.pone.0225693> PMID: 31794567
25. Santos-Peñate DR, Campos-Rodríguez CM, Moreno-Pérez JA. A kernel search matheuristic to solve the discrete leader-follower location problem. *Networks and Spatial Economics*. 2020; 20:73–98. <https://doi.org/10.1007/s11067-019-09472-7>
26. Saha AK, Paul A, Azeem A, Paul SK. Mitigating partial-disruption risk: A joint facility location and inventory model considering customers' preferences and the role of substitute products and backorder offers. *Computers & Operations Research*. 2020; 117:104884. <https://doi.org/10.1016/j.cor.2020.104884>
27. Yu Z, Khan SAR. Evolutionary game analysis of green agricultural product supply chain financing system: COVID-19 pandemic. *International Journal of Logistics Research and Applications*. 2022; 25:1115–1135. <https://doi.org/10.1080/13675567.2021.1879752>
28. Lin YH, Tian Q, Zhao Y. Locating facilities under competition and market expansion: Formulation, optimization, and implications. *Production and Operations Management*. 2022. <https://doi.org/10.1111/poms.13737>
29. Latifi SE, Tavakkoli-Moghaddam R, Fazeli E, Arefkhani H. Competitive facility location problem with foresight considering discrete-nature attractiveness for facilities: Model and solution. *Computers & Operations Research*. 2022; 146:105900. <https://doi.org/10.1016/j.cor.2022.105900>
30. Martello S, Toth P. *Knapsack problems: algorithms and computer implementations*. John Wiley & Sons, 1990.
31. Yu W. Robust model for discrete competitive facility location problem with the uncertainty of customer behaviors. *Optimization Letters*. 2020; 14:2107–2125. <https://doi.org/10.1007/s11590-020-01538-7>
32. Beasley JE. OR-library: Distributing test problems by electronic mail. *Journal of the Operational Research Society*. 1990; 41:1069–1072. <https://doi.org/10.2307/2582903>
33. Daskin MS. *Network and discrete location: models, algorithms, and applications*. John Wiley & Sons, 1995.