No. 021/2020/POM/DEC

A single-facility competitive location problem in the plane based on customer choice rules

Hongguang Ma

School of Economics and Management

Beijing University of Chemical Technology

Xiang Li*

School of Economics and Management

Beijing University of Chemical Technology

Xiaoyu Guan

School of Economics and Management

Beijing University of Chemical Technology

Xiande Zhao

Department of Economics and Decision Sciences

China Europe International Business School (CEIBS)

Liang Wang

Department of Economics and Decision Sciences

China Europe International Business School (CEIBS)

June 2020

* Corresponding author: Xiang Li (lixiang@mail.buct.edu.cn). Address: School of Economics and Management, Beijing University of Chemical Technology, 15 Beisanhuan East Road, Chaoyang, Beijing 100029, China. The authors would like to thank financial support from National Natural Science Foundation of China (Nos. 71722007 & 71931001), the Funds for First-class Discipline Construction (XK1802-5) and the Fundamental Research Funds for the Central Universities (buctrc201926).

A single-facility competitive location problem in the plane based on customer choice rules

Hongguang Ma[a], Xiang Li[a,*], Xiaoyu Guan[a], Xiande Zhao[b], Liang Wang[b]

[a]*School of Economics and Management, Beijing University of Chemical Technology, Beijing 100029, China*
[b]*China Europe International Business School, 699 Hongfeng Road, Pudong, Shanghai 201206, China*

Abstract

Since customer choice rules would greatly affect the performance of retail facilities, they should be considered when a chain wants to locate a new facility in a competitive market. In the existing studies, customers' choice behavior is usually considered as homogeneous, which means that all customers patronize facilities with one kind of customer choice rules: the deterministic rule, the probabilistic rule or the multi-deterministic rule. However, it is not in line with reality as we have investigated people's choice behavior on convenience stores by questionnaire surveys, and survey results show that different customers may patronize facilities with different choice rules. In order to study competitive facility location problems in which customers' choice behavior is heterogeneous, we classify customers as three types by customer choice rules, the relative proportions of which are calculated based on questionnaires. A customer classification based competitive facility location model in the plane is proposed in which location and quality of the new facility are to be determined in order to maximize the profit of the locating chain. Since the model is non-convex and discontinuous, and location problems in practice are usually large-scale, four kinds of heuristic algorithms instead of exact algorithms are designed for obtaining a satisfactory solution including Particle Swarm Optimization, Tabu Search, Simulated Annealing and Genetic Algorithm. Numerical experiments show that Particle Swarm Optimization performs best both in computation efficiency and solution precision. Comparisons among location results employing different customer proportions reveal that customer proportion significantly affects location results. Most importantly, the locating chain may lose large profit once the customer proportion is wrongly estimated. Maximum profit loss is more than 20% in our cases.

Keywords: Competitive facility location, Customer choice rules, Heuristic algorithms, Customer classification

*Corresponding author.
Email address: lixiang@mail.buct.edu.cn (Xiang Li)

1. Introduction

How to locate a facility in a competitive market environment has become a major concern for managers. Where to locate a new facility is a vitally important strategic decision for a firm, as the location has a major and long-term influence on the profit of the firm, and it cannot be easily relocated as setting up a facility usually requires a massive investment. Taking convenience store industry as an example, convenience stores develop in China since 1990s. Chains are constantly working on finding good locations to open new stores in order to obtain larger market share and profits. The number of convenience stores in China has exceeded 122 thousands up to 2018, maintaining a growth rate of over 10 percent in recent years. In the meantime, many convenience stores close the door because of poor location. In 2018, Lin Jia, a famous convenience store chain, closed its 168 stores in Beijing. In the increasingly drastic market, location of convenience stores is particularly important.

Many researchers have been focusing on the competitive facility location problem and a number of competitive location models can be found in the literature. Considering the competitive facility location is a problem of wide range which can be applied to many practical cases, competitive location models vary in the ingredients which form the models. Some competitive location models are summarized in Table 1.

Table 1: Summary of competitive location models.

Location space	The plane (Fernández et al., 2019)	A discrete set (Fernández et al., 2017b)
The number of new facilities	Single (Blanquero et al., 2014)	Multiple (Blanquero et al., 2016)
The market environment	Static (Plastria, 2001)	Dynamic (Drezner, 2009)
The demand	Elastic (Redondo et al., 2012)	Inelastic (Redondo et al., 2009)

In all of these location studies, one of the most important aspects for making a better location decision is to understand how customers make choice among all facilities in an area. Various approaches are proposed to formulate the interrelationship among customers and facilities in the literature, including the proximity approach, the utility approach, the cover-based approach and the gravity model, etc. (Drezner, 2014). With the development of different approaches, different kinds of customer choice rules are proposed. The commonly used customer choice rules in the literature can be classified into three kinds:

- The deterministic rule: A customer only patronizes the facility that attracted him/her most (Drezner, 1994a).
- The probabilistic rule: A customer splits his/her demand among all facilities in an area proportionally to the attraction of each facility to him/her (Drezner, 1994b).

- The multi-deterministic rule: A customer split his/her demand among chains by patronising only one facility from each chain, the one with the highest attraction, and the demand is split among those facilities proportionally to their attraction (Fernández et al., 2017a).

Although each of these customer choice rules has been used to calculate the market share captured by facilities and good estimations are obtained in some cases, they still have trouble representing customer behavior under more general conditions. Existing studies assume customers patronize facilities with one kind of choice rules, but what is closer to reality is that different customers may patronize based on different rules. Customers who prefer to patronize the nearest facility or have a strong preference for one particular facility are more likely to adopt the deterministic rule. Customers who have no obvious preference among facilities may patronize as the probabilistic rule. The multi-deterministic rule is for those customers who would go to different facilities depending on the goods he/she would buy. Traditional location studies mostly focus on the location of generalized facilities instead of a specific industry. However, customer behavior in different industries has quite large difference. For instance, compared to snacks, people show more loyalty when they buy electronic products, leading to a less proportion of customers using the probabilistic rule. Customers are more likely to behave with different choice rules when they patronize a convenience store as convenience stores are more and more widespread and the goods they sold are basically homogeneous.

We analysed the choice behavior of 123 people near Beijing University of Chemistry Technology choosing convenience stores through questionnaire surveys. In the questionnaire, we listed three options corresponding to three kinds of customer choice rules. The result of customer choice rules are shown in Table 2. It reveals that these three customer choice rules are able to represent most people's choice behavior on convenience stores. Questionnaire results show that each of these three choice rules is employed by a part of people when they choose convenience stores to patronize. Table 2 shows that the customer proportion of the deterministic type, the probabilistic type and the multi-deterministic type on convenience stores is 30.33%: 40.16%: 29.51%.

Table 2: The questionnaire results.

Customer type	The number of customers	Customer proportion
The deterministic type	37	30.08%
The probabilistic type	49	39.84%
The multi-deterministic type	36	29.51%
Others	1	0.01%

In this article, we classify customers into three types according to survey results rather than one type. Furthermore, a customer classification based facility location model is formulated in which the customer proportion (the relative proportions of three types' customers) is considered. The model constructed in this way is closer to reality and has a better estimation of the market share captured by each facility. The main contributions of this paper are as follows:

(1) In order to formulate customers' choice behavior as realistic as possible, we classify customers into three types according to the questionnaire surveys for investigating customers' choice behavior on convenience stores. A customer classification based facility location model in the plane is formulated. Location and quality of the new facility are to be determined in order to maximize the profit of the locating chain.
(2) Four heuristic algorithms, including: Particle Swarm Optimization (PSO), Tabu Search (TS), Simulated Annealing (SA) and Genetic Algorithm (GA), are designed to solve this model. We compared the performance of these four heuristic algorithms on solving our problem through numerical experiments and results show that PSO is superior to other three algorithms in computation efficiency and solution precision.
(3) The influence of different customer proportions on the location results is investigated. In particular, a set of location problems are generated to calculate the profit loss of the locating chain when the customer proportion is wrongly employed.

The remainder of this paper is organized as follows. In Section 2, we review some important literature related to our work. Section 3 formulates a single facility location model in which the customers are classified into three types. Four heuristic algorithms are designed to solve the model in Section 4. Section 5 presents some computational experiments for evaluating the influence of employing different proportions of customers on location results and comparing the performance of these algorithms. Finally, Section 6 summarizes the paper briefly and gives some directions for future research.

2. Literature review

Existing facility location studies provide various location models to address different kinds of location cases. We first review the competitive location problems and relevant approaches for modeling the competition relationship. Furthermore, we introduce three customer choice rules which are commonly employed in location problems. Finally, relevant methods for solving location models are reviewed.

2.1. Competitive location problems and relevant approaches

Considering the market in the reality is competitive generally, most facility location studies aim at solving competitive facility location problems. The competitive location problem was first proposed by Hotelling (1990). He discussed the location problem of two competitive sellers locating one facility each on a linear market, and customers choose one of two facilities to patronize. From then on a surge of studies in competitive location arise. For an overview of competitive problems, readers could refer to Plastria (2001), Eiselt and Laporte (1997), Drezner (2014). In the competitive location studies, various approaches have been developed to model the competitive relationship of facilities, which are the basis for estimating the market share captured by facilities. The proximity approach, which assumed customers patronize the closest facility, was first used in Hotelling (1990). Drezner (1994a) developed the utility approach in which customers consider the utility of facilities rather than just the distance. Drezner et al. (2011) introduced the cover-based approach to estimating market share. Each competing facility has a coverage area, the size of which depends on the attractiveness of the facility. A customer in a facility's coverage area is attracted to the facility. Both the utility approach and the cover-based approach imply an important assumption: the "all or nothing" assumption, which means for a certain demand point, a facility captures either all the market share of the demand point, or nothing. Except these approaches, another more commonly used approach for estimating market share is the gravity model. Most competitive location problems in the plane have recently used the gravity model (Drezner, 2014).

The gravity model is based on Newton's law of universal gravitation, and has been used as the basic of numerous facility location models. Huff (1964) formulated the probability of a customer at a given demand point traveling to a shopping center as a function of attraction he/she feels from shopping centers in the area. In the gravity model, the attraction of a facility for a customer is modeled using a function which is non-decreasing with respect to quality of the facility and non-increasing with respect to the distance between the facility and the customer. In Huff's model, the quality of a facility is represented by the size (floor area) of the shopping center, which is also described as its service level in some other studies (Dan and Marcotte, 2019). Drezner (1994b) formulated a single retail facility location model in the plane based on the gravity model. The objective of the model is to find the optimal location of a new facility so as to maximize the market share captured by the locating chain with given quality value of the new facility. Rather than "all or nothing" assumption, the gravity model formulates the attraction of facilities to demand points, based on which customers' varied behavior can be interpreted. The gravity model has been successfully applied in many competitive location studies due to its great merits such as simplicity and ease of understanding (Fernández et al., 2017a; Blanquero et al., 2016; Drezner and Drezner, 2004). Thus we also use the gravity model to formulate the location model in our study.

2.2. Customer choice rules

Generally, the purpose of competitive facility location is to obtain higher market share or profit of the locating chain. In order to estimate the market share properly, we need to understand how customers choose facilities to patronize. Customers' choice behaviour is an important factor that must be taken into account when chains make location decisions. According to existing location studies, customers' choice behavior can be summarized into three kinds of rules.

The first one is the deterministic rule which assumes that a customer only patronizes the facility that attracted him/her most. The deterministic rule is also called binary rule, first proposed in Hotelling (1990) for a location problem in the line market. Later, this customer choice rule has been widely used in many competitive facility location models, especially in the $(r|X_p)$-medianoid model which was originally introduced by Hakimi (1983).

The second customer choice rule commonly used in the literature is the probabilistic rule which assumes that a customer splits his/her demand over all facilities in an area proportionally to the attraction of these facilities. It is also known as the proportional rule, first investigated by Huff (1964), along with the gravity model. The probabilistic rule has been widely used in the location studies especially in Huff-like location problems (Drezner, 2009; Drezner and Drezner, 2004; Drezner, 1998).

The third customer choice rule is the multi-deterministic choice rule suggested by Hakimi (1990), also named as the partially binary rule. It assumes that a customer split his/her demand among the most attractive facilities of each chain, and the demand is split among those facilities proportionally to their attraction. The multi-deterministic choice rule is different from the probabilistic rule as customers with this rule won't patronize all facilities that belong to the same chain. It is also distinct from the deterministic rule as customers following this rule patronize more than one single facility. Following this idea, Serra and Colomé (2001) and Fernández et al. (2017a) studied location problems in discrete points and in the plane, respectively.

The location problems employing different kinds of customer choice rules has been considered in some literature. For example, Suárez-Vega et al. (2004) studied continuous location problems with six scenarios considering three customer choice rules and two types of goods (essential or unessential). In each scenario one kind of customer choice rule and one type of goods are considered. Biesinger et al. (2016) also employed the same scenarios when formulating location models on discrete space.

Based on the above literature, customers' choice behavior is an inevitable issue when a chain wants to locate a new facility. Previous researchers mainly assume all customers patronize with the same kind of choice rule. However, customers may patronize with different choice rules according to our survey for investigating customers' choice behavior on convenience stores. To the extent of our knowledge, there is no research which has studied the location problem with different types of customers. In this paper, we classify customers into three types by customer choice rules and investigate the influence of different customer proportions on the location results and on the profit obtained by the locating chain.

2.3. Methods for solving location problems

Since many competitive location models are difficult to solve because of the non-convexity and discontinuity of the objective function, heuristic algorithms are widely used to obtain acceptable solutions in location studies. A multi-start heuristic algorithm, the Weiszfeld-like algorithm, was proposed in Fernández et al. (2007) for solving the location problem in the plane. Redondo et al. (2009) investigated an evolutionary algorithm called UEGO for solving the competitive facility location problem with a leader and a follower. Genetic Algorithm was designed for solving discrete competitive facility location problems in Lančinskas et al. (2017). Tabu Search was employed to solve a bi-level facility location model in Shan et al. (2019). A two-phase heuristic approach was designed for the discrete dynamic location problem in da Gama and Captivo (1998). All these heuristic algorithms showed good performance in solving corresponding location problems. Three heuristic algorithms, including Tabu Search (TS), Simulated Annealing (SA) and Genetic Algorithm (GA), were designed to solve facility location problems in Arostegui Jr et al. (2006). The comparison results indicated that Tabu Search performed best in most cases. In this study, we design PSO, GA, SA and TS for our problem and compare the relative performance of these algorithms in solving our location model.

3. Problem description

We consider a competitive single facility location problem with the consideration of customer classification. The demand, assumed to be inelastic, is assembled at demand points. Locations and demand volume of demand points are given. Locations and the quality of existing facilities are also known. Considering the location problem we are studying is continuous, we use Euclidean distance formula to calculate the distance between two places in our problem. The attraction of facility j that a customer at demand point i feels is formulated as quality of the facility divided by the distance effect: $u_{ij} = \alpha_j / g_i(d_{ij})$. Figure 1 is drawn to obtain an intuitive image of the location problem. In Figure 1, 20 demand points denoted by dots distribute in a 10 by 10 miles area. There are 2 competing chains in this area, owning 3 and 2 existing facilities respectively, represented by triangles and squares. Different graphics are used to represent facilities belonging to different chains. The size of a triangle or a square is construed as the quality of the corresponding facility. The bigger a triangle or a square is, the higher the quality of the corresponding facility is. Similarly, the size of a dot denotes the demand volume of the corresponding demand point. Now chain 1 is about to locate a new facility in the area.

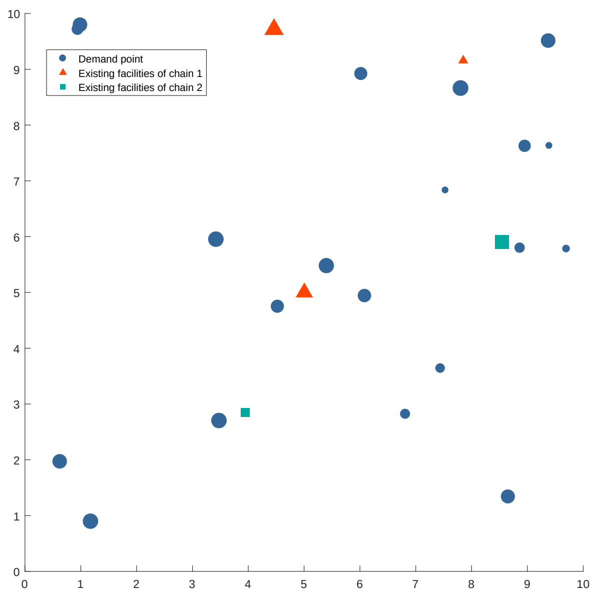

Figure 1: Problem description.

Customers are classified into three kinds by three customer choice rules. The deterministic customers patronize the facility which has the largest attraction to them. The probabilistic customers probabilistically patronize all the facilities in the area in proportion to the attraction of these facilities. The multi-deterministic customers probabilistically patronize the most attractive facility of each chain proportionally to the attraction of these facilities. The relative proportions of these three types' customers have been obtained through questionnaires.

As customer behavior is highly subjective, the three kinds of customer choice rule may not be able to reflect the choice behavior of every customer. Some customers may choose facilities by different choice rules in different situations. A customer could be the deterministic customer sometimes, and be the probabilistic customer at other times. This kinds of uncertainty makes the customer proportion uncertain, and further makes the estimated market share uncertain. For obtaining better location result, our location model should be as close to the reality as possible. So we consider that a demand point may have multiple possible customer proportion. Based on each kind of proportion the demand point's market share that the locating chain obtains can be calculated. Then the demand point's market share that the locating chain obtains can be represented as a fuzzy variable. The possibility of each element in the fuzzy variable is explained as the closeness of corresponding proportion to the actual customer proportion in the demand point. For the convenience of computing, we calculate the expected value of fuzzy market share, and formulate the location model based on the expected value.

The purpose of this problem is finding the best place to locate the new facility and determining the optimal quality in order to maximize the profit obtained by the locating chain. The profit of the chain is calculated as the income due to the market share minus the operational costs.

3.1. Notations

For better understanding the mathematical formulations in this paper, notations that will be used are shown in Table 3.

Table 3: List of notations.

Notations	
Sets and indexes	
n	number of demand points
m	number of existing facilities
k	number of existing chains
i	index of demand points, $i = 1, ..., n$
c	index of existing chains, $c = 1, ..., k$ (chain 1 is the locating chain)
j	index of existing facilities, $j = 1, ..., m$ (we assume that from $j = j^1_{\min}(= 1)$ to $j^1_{\max}$ the facilities belong to chain $c = 1(j^1_{\min} < j^1_{\max})$; from $j = j^2_{\min}(= j^1_{\max} + 1)$ to $j^2_{\max}$ belong to chain $c = 2$;...; from $j = j^k_{\min}(= j^{k-1}_{\max} + 1)$ to $j^k_{\max}(= m)$ belong to chain $c = k$)

Decision variables

x	location of the new facility, $x = (x_1, x_2)$
α	quality of the new facility

Parameters

p_i	location of demand point i, $p_i = (p_{i1}, p_{i2})$
w_i	demand volume (or buying power) at p_i, $w_i > 0$
f_j	location of existing facility j, $f_j = (f_{j1}, f_{j2})$
d_{ij}	distance between demand point i and facility j, $d_{ij} > 0$, $d_{ij} = \sqrt{(p_{i1} - f_{j1})^2 + (p_{i2} - f_{j2})^2}$
$d_i(x)$	distance between demand point i and the new facility located at x
α_j	quality of facility j, $\alpha_j > 0$
$\alpha_{\min}$	minimum level of quality for the new facility, $\alpha_{\min} > 0$
$\alpha_{\max}$	maximum level of quality for the new facility, $\alpha_{\min} \geq \alpha_{\max}$
S	region of the plane where the new facility can be located
$g(\cdot)$	a continuous non-negative and non-decreasing function, which modulates the decrease of attractiveness as a function of distance
u_{ij}	attraction that demand point i feels for the facility$j$ (or utility of the facility j perceived by people at demand point i). In this paper, $u_{ij} = \alpha_j / g_i(d_{ij})$
$u_i(x, \alpha)$	attraction that customers in demand point i feel for the new facility located at x with quality α
u_i^c	maximum attraction that customers in demand point i feel for all of the existing facilities of chain c, $u_i^c = \max\{u_{ij} : j = j^c_{\min}, ..., j^c_{\max}\}$
λ_1	the proportion of customers using the deterministic rule
λ_2	the proportion of customers using the probabilistic rule
$A(x, \alpha)$	set of demand points to which the maximum attraction of the locating chain's facilities (including chain 1's existing facilities and the new facility located at x with quality α) is larger than other chains' facilities, $A(x, \alpha) = \{i \vert \max\{u_i^1, u_i(x, \alpha)\} \geq \max\{u_i^c : c = 2, ..., k\}\}$
$M_D(x, \alpha)$	the market share captured by the chain 1's existing facilities and by the new facility located at x with quality α when customers are deterministic
$M_P(x, \alpha)$	the market share captured by the chain 1's existing facilities and by the new facility located at x with quality α when customers are probabilistic
$M_M(x, \alpha)$	the market share captured by the chain 1's existing facilities and by the new facility located at x with quality α when customers are multi-deterministic
$M_C(x, \alpha)$	the market share captured by the chain 1's existing facilities and by the new facility located at x with quality α when customers are classified into three types by customer choice rules
$F(\cdot)$	a function that converts the market share into expected sales
$G(x, \alpha)$	a function for evaluating the operational costs of a facility located at x with quality α

E	the operational costs of chain 1's existing facilities

3.2. Market share obtained by employing single customer choice rule

We first formulate the market share obtained by the locating chain when customers' choice behavior is considered as homogeneous, which means that all customers patronize facilities with one kind of customer choice rules.

In the deterministic rule, it is assumed that a customer at a demand point will only patronize the facility that has the largest attraction to the customer, disregarding other facilities which are less attractive. We define $A(x, \alpha)$ as the set of demand points to which the maximum attraction of the locating chain's facilities is larger than other chains' facilities. When all customers in demand i are the deterministic type, the market share of the demand point captured by the chain 1 is

$$w_{Di}(x, \alpha) = \begin{cases} 0 & i \notin A(x, \alpha) \\ w_i & i \in A(x, \alpha) \end{cases} \tag{1}$$

where w_i denotes the demand volume at demand point i. Then the market share of all demand points captured by the chain 1 is

$$M_D(x, \alpha) = \sum_{i=1}^{n} w_{Di}(x, \alpha). \tag{2}$$

In the probabilistic rule, it is assumed that that a customer splits his/her demand over all facilities in an area proportionally to the attraction he/she feels from each facility. When all customers in demand i are the probabilistic type, the market share captured by the chain 1 is

$$w_{Pi}(x, \alpha) = w_i \times \frac{u_i(x, \alpha) + \sum_{j=1}^{j^1_{\max}} u_{ij}}{u_i(x, \alpha) + \sum_{j=1}^{m} u_{ij}}, \tag{3}$$

where u_{ij} is the attraction that demand point i feels for the facility$j$, $u_{ij} = \alpha_j / g_i(d_{ij})$, and $u_i(x, \alpha)$ is the attraction that customers in demand point i feel for the new facility. The function $g(d)$ in the attraction function is a continuous non-negative and non-decreasing function. The form of $g(d)$ is usually defined as $g(d) = d^\eta$ for some $\eta > 0$ (Huff, 1964; Nakanishi and Cooper, 1974). Then the market share of all demand points captured by the chain 1 is

$$M_P(x,\alpha) = \sum_{i=1}^{n} w_{Pi}(x,\alpha). \tag{4}$$

In the multi-deterministic rule, it is assumed that a customer probabilistically patronizes the most attractive facilities from each chain in the area. He/she will patronize these facilities proportionally to the attraction he/she feels from each facility. When all customers in demand i are the multi-deterministic type, the market share captured by the chain 1 is

$$w_{Mi}(x,\alpha) = w_i \times \frac{\max\{u_i(x,\alpha), u_i^1\}}{\max\{u_i(x,\alpha), u_i^1\} + \sum_{c=2}^{k} u_i^c}, \tag{5}$$

where u_i^c is the maximum attraction that customers in demand point i feel for all of the existing facilities of chain c, $u_i^c = \max\{u_{ij} : j = j_{\min}^c, ..., j_{\max}^c\}$.Then the market share of all demand points captured by the chain 1 is

$$M_M(x,\alpha) = \sum_{i=1}^{n} w_{Mi}(x,\alpha). \tag{6}$$

3.3. Fuzzy market share obtained by customer classification

Next we formulate the market share obtained by the locating chain when customers' choice behavior is considered as heterogeneous, which means that different customers patronize facilities with different kinds of customer choice rules. Customers are classified into three kinds by customer choice rules. The market share of chain 1 is estimated based on the customer classification.

However, in the reality, some customers' choice behavior may be uncertain. They may patronize facilities by different choice rules in different situations. For example, a customer may mostly patronize as the deterministic rule, and sometimes patronizes as the probabilistic rule. For these customers, one kind of choice rule can not accurately reflect their choice behavior. A customer could be the deterministic type sometimes, and be the probabilistic type at other times. A demand point may have multiple possible customer proportions. This uncertainty makes the customer proportion uncertain, and further makes the market share which is calculated based on the customer proportion uncertain. For better estimating the market share, we use fuzzy variables instead of real variables to estimate each demand point's market share captured by chain 1.

For each demand point i, there are s_i kinds of possible customer proportions:

$$\begin{cases} b_i^1 = \lambda_{i1}^1 : \lambda_{i2}^1 : \lambda_{i3}^1 \quad with \quad \gamma_i(b_i^1) = \mu_i^1, \\ b_i^2 = \lambda_{i1}^2 : \lambda_{i2}^2 : \lambda_{i3}^2 \quad with \quad \gamma_i(b_i^2) = \mu_i^2, \\ ..., \\ b_i^{s_i} = \lambda_{i1}^{s_i} : \lambda_{i2}^{s_i} : \lambda_{i3}^{s_i} \quad with \quad \gamma_i(b_i^{s_i}) = \mu_i^{s_i}. \end{cases} \tag{7}$$

Based on these customer proportions, demand point i's estimated market share obtained by chain 1 can be calculated:

$$\begin{cases} w_{Fi}^{1}(x,\alpha) = \lambda_{i1}^{1} \times w_{Di}(x,\alpha) + \lambda_{i2}^{1} \times w_{Pi}(x,\alpha) + \lambda_{i3}^{1} \times w_{Mi}(x,\alpha), \\ w_{Fi}^{2}(x,\alpha) = \lambda_{i1}^{2} \times w_{Di}(x,\alpha) + \lambda_{i2}^{2} \times w_{Pi}(x,\alpha) + \lambda_{i3}^{2} \times w_{Mi}(x,\alpha), \\ \dots, \\ w_{Fi}^{s_i}(x,\alpha) = \lambda_{i1}^{s_i} \times w_{Di}(x,\alpha) + \lambda_{i2}^{s_i} \times w_{Pi}(x,\alpha) + \lambda_{i3}^{s_i} \times w_{Mi}(x,\alpha). \end{cases} \tag{8}$$

which can form a fuzzy variable:

$$\tilde{w}_{Fi}(x,\alpha) = \frac{\gamma_i(b_i^1)}{w_{Fi}^{1}(x,\alpha)} + \frac{\gamma_i(b_i^2)}{w_{Fi}^{2}(x,\alpha)} + \dots + \frac{\gamma_i(b_i^{s_i})}{w_{Fi}^{s_i}(x,\alpha)}. \tag{9}$$

As different demand points may have different numbers of customer proportions, the structure of fuzzy variables which are used to represent the market share of different demand points may be different. To facilitate easy calculation, we calculate the mathematical expectation of $\tilde{w_{Fi}}(x,\alpha)$. We follow the method proposed by Liu and Liu (2002) to calculate the mathematical expectation of fuzzy variables:

$$E(\tilde{w}_{Fi}(x,\alpha)) = \sum_{l=1}^{s_i} (e_i^l \times w_{Fi}^{1}(x,\alpha)) \tag{10}$$

where

$$e_i^l = \frac{1}{2} \times (\max_{1\le q\le l} \mu_i^1 - \max_{0\le q\le l-1} \mu_i^1) + \frac{1}{2} \times (\max_{l\le q\le s_i} \mu_i^1 - \max_{l+1\le q\le s_i+1} \mu_i^1), \mu_i^0 = \mu_i^{s_i+1} = 0. \tag{11}$$

Then the market share of all demand points captured by the chain 1 is

$$M_F(x,\alpha) = \sum_{i=1}^{n} E(\tilde{w}_{Fi}(x,\alpha)). \tag{12}$$

3.4. Profit maximization model

The aim of our problem is to maximize the profit of the locating chain. The problem to be solved is then

$$\begin{cases} \max & \Pi(x,\alpha) = F(M_F(x,\alpha)) - G(x,\alpha) - E \\ \text{s. t.} & \alpha \in [\alpha_{\min}, \alpha_{\max}] \\ & x \in S \end{cases} \tag{13}$$

where $F(\cdot)$ is a strictly increasing real valued function which converts the market share into expected sales, $G(x,\alpha)$ is a function for evaluating the operational costs of the new facility when the new facility is located at x with quality α, $\Pi(x,\alpha)$ is the profit obtained by the chain, and E is the operational costs of the locating chain's existing facilities. The constraint $\alpha \in [\alpha_{\min}, \alpha_{\max}]$ gives the minimum value and the maximum value that the quality of a facility may take in practice. The region of the plane where the new facility can be located is denoted by S. The function F is always assumed to be linear, for example, $F(M_C(x,\alpha)) = c \cdot M_C(x,\alpha)$, where c is the average income per unit of good sold (Fernández et al., 2007). The operational cost function $G(x,\alpha)$ is considered to be separable, i.e., of the form $G(x,\alpha) = G_1(x) + G_2(\alpha)$, as it is influenced by the variables x and α respectively. The function $G_1(x)$ increases when x is closer to demand points, which reflects the higher operational costs when the facility is located around the demand points. The function $G_1(x)$ also increases when w_i is bigger, since the operational costs is higher when the facility is located around places with more people. Possible expressions for $G_1(x)$ which can be found in Fernández et al. (2007) are

$$G_1(x) = \sum_{i=1}^{n} \phi_i(d_{ix}),$$

with

$$\phi_i(d_{ix}) = w_i/((d_i(x))^{\phi_{i0}} + \phi_{i1}), \quad \phi_{i0},\, \phi_{i1} > 0,$$

or

$$\phi_i(d_{ix}) = w_i/(e^{\frac{d_i(x)}{\phi_{i0}}} - 1 + \phi_{i1}), \quad \phi_{i0},\, \phi_{i1} > 0.$$

The function $G_2(\alpha)$ should be a non-decreasing and convex function, since the higher the quality of the facility is, the higher the operational costs is, at an increasing rate. A few typical forms that can be found in Fernández et al. (2007) for $G_2(\alpha)$ are

$$G_2(\alpha) = (\alpha/\alpha_0)^{\alpha_1}, \quad \alpha_0 > 0,\, \alpha_1 \geq 1,$$

or

$$G_2(\alpha) = e^{\frac{\alpha}{\alpha_0} + \alpha_1} - e^{\alpha_1}, \quad \alpha_0 > 0.$$

4. Algorithm

With the abundance of location research, many methods are employed to solve location models. Drezner and Drezner (2004) devise a the Big Triangle Small Triangle method to solve the location problem in which the decision variable is the location of the new facility and customers are the probabilistic type. In our location model, the decision variables are location and quality of the new facility, and customers are three different type, which makes the objective function discontinuous. These characteristics makes the Big Triangle Small Triangle method unsuitable for our model. As most of competitive facility location models are non-convex and discontinuous, heuristic algorithms have been widely used in the location studies. An empirical comparison of Tabu Search (TS), Simulated Annealing (SA), and Genetic Algorithm (GA) for facility location problems can be found in Arostegui Jr et al. (2006). In Özgün-Kibiroğlu et al. (2019), Particle Swarm optimization (PSO) was designed to solve a location model and good solutions are obtained. In this section, we also design PSO for our problem, and compare the performance of PSO, GA, TS and SA in solving our location model.

4.1. Particle Swarm Optimization

Particle Swarm Optimization (PSO) is a novel random optimization method with simple structure, easy realization and good utility. It is based on swarm intelligence which has a powerful ability of global optimization. What is most important is that the implementation procedure of PSO has less requirement on the objective function. Considering the function constructed in this paper is non-convex and discontinuous, PSO is an appropriate heuristic algorithm to solve this problem.

Particle Swarm Optimization Algorithm was first proposed by Eberhart and Kennedy (1995). It is a swarm intelligence based evolutionary algorithm inspired by the regularity of birds' movement. When birds are searching food in a space, they have no idea where the food is, but they know how far they are from the food roughly, so the best search strategy is searching around the bird which is nearest to the food. PSO's principle is similar to the strategy of birds searching food. Every bird is called a particle in the algorithm with two properties: position and speed. Every position in the solution space represents a feasible solution. Speed of a particle determines its next move's direction and distance. A fitness function is defined to judge if a position is good or bad. In each iteration every particle's speed is calculated based on the messages of their own historical best position and the total swarm's current best position. Particles will be clustered in the area where is near to the optimal solution gradually. Then they keep searching around the area and the optimal solution or the near optimal solution will be found.

We first randomly generate Q particles to constitute a swarm. As there are 3 values (longitude, latitude and quality value of the new facility) to be determined in our problem, each particle's position is generated as a 3 dimensional vector, for example, particle q's position $X_q = (x_{q1}, x_{q2}, \alpha_q), q = 1, 2, ..., Q$. Its speed is also a 3 dimensional vector $V_q = (v_{q1}, v_{q2}, v_{q3}), q = 1, 2, ..., Q$. The objective $\Pi(X_q)$ is set as the fitness function in our algorithm, which is used to measure if a position is good or bad. The personal best position that particle q has found in the searching process is denoted as $Pbest_q = (pb_{q1}, pb_{q2}, pb_{q3}), q = 1, 2, ..., Q$. The global best position that the swarm has found in the searching process is denoted as $Gbest = (pg_1, pg_2, pg_3)$. After each iteration, these two positions can be obtained, and particles will update their positions and speeds as

$$\begin{cases} v_{q1} = h \times v_{q1} + c_1 \times r_1 \times (pb_{q1} - x_{q1}) + c_2 \times r_2 \times (pg_1 - x_{q1}) \\ v_{q2} = h \times v_{q2} + c_1 \times r_1 \times (pb_{q2} - x_{q2}) + c_2 \times r_2 \times (pg_2 - x_{q2}) \\ v_{q3} = h \times v_{q3} + c_1 \times r_1 \times (pb_{q3} - \alpha_q) + c_2 \times r_2 \times (pg_3 - \alpha_q) \\ x_{q1} = x_{q1} + v_{q1} \\ x_{q2} = x_{q2} + v_{q2} \\ \alpha_q = \alpha_q + v_{q3}, \quad 1 \leq q \leq Q, \end{cases} \tag{14}$$

where h is the inertia coefficient that adjusts the balance between the global search and the local search. A higher value of h will improve the algorithm's global search ability, and a lower value of h will increase the algorithm's local search ability. Higher global ability is required at the algorithm's early stage so as to enlarge the search scopes. Higher local search ability is better for the algorithm's later period so that the algorithm can rapidly converge to optimum. In our research, the inertia coefficient h is dynamically adjusted as follow

$$h = \frac{e^{-k} \times (h_{max} - h_{min}) \times (k_{max} - k)}{k_{max}} + h_{min}, \tag{15}$$

where h_{max} and h_{min} are the maximum value and the minimum value of the inertia coefficient; k is the current generation number; k_{max} is the maximum generation number.

r_1 and r_2 are two random numbers distributed in the range (0,1); c_1 and c_2 are acceleration coefficients representing the particle's dependence on itself (self cognition behavior) and on the swarm (social behavior), respectively. A relatively high value of c_1 will encourage particles to move toward their local best positions, while higher values of c_2 will result in faster convergence to the swarm best position. In practice the value of c_1 are usually set equal to the value of c_2 in order to represents an equal emphasis on both directions. So c_1 and c_2 are called acceleration coefficients below. The process of Particle Swarm Optimization used in this article is summarized below.

Step 1: Initialize the particle swarm: determine the swarm size Q, randomly generate the position X_q and the speed V_q of each particle.

Step 2: Calculate the fitness value $\Pi(X_q)$ of each particle.

Step 3: For each particle, compare its fitness value $\Pi(X_q)$ with its personal best value $\Pi(Pbest_q)$; if $\Pi(X_q) > \Pi(Pbest_q)$, then update its personal best position $Pbest_q$, $Pbest_q = X_q$.

Step 4: Find the maximum personal best value among all particles; assign the corresponding position value to the global best position $Gbest$.

Step 5: Update particles'positions and speed according to the formulation 14.

Step 6: If the terminal condition (the maximum number of iterations has been reached) is met, end the iteration and output the global best value and the corresponding position; otherwise back to Step 2.

4.2. Genetic Algorithm

Genetic Algorithm (GA) is an evolutionary algorithm which is derived from the evolutionary law of biology in the natural environment. GA was first proposed by Holland (1992) and has becoming one of the most popular heuristic algorithms for optimization problems. Starting from an initial population, the algorithm evaluates the fitness of each individual in the population. Then the individual with a higher fitness value could be selected with a higher probability to produce offspring after crossover and mutation process. After several generations the individual with higher fitness value is generated. In our paper, we adopt binary codes to design Genetic Algorithm. As the problem we studied is to locate a new facility in the continuous space, our decision variables are continuous variables. In order to use binary codes to represent solutions, we first need to calculate how many binary numbers is required. For example, the ranges of new facility's abscissa value x_1 and ordinate value x_2 are from 0 to 10, and the range of new facility's quality value α is from 0.5 to 5. If x_1 is accurate to 3 decimal places, at least we need $10^4 (= 10 * 10^3)$ numbers. Since $2^{13} < 10^4 < 2^{14}$, we need 14 binary numbers to represent a solution of x_1. By this way, we can calculate that we need 41 binary numbers to form a chromosome as a solution including x_1, x_2 and α. When we calculate the fitness value of each chromosome, we first need to convert binary to decimal, then convert decimal to the numbers of our range. (15) and (16) gives an example of converting a fourteen-digit binary number to the range of [0, 10]:

$$(b_0...b_{12}b_{13})_2 = (\sum_{i=0}^{13}(b_i \times 2^i))_{10} = x^{'}. \tag{16}$$

$$x = 0 + x^{'} \times \frac{10-0}{2^{14}-1}. \tag{17}$$

For each generation, we choose the chromosomes with higher fitness value to generate the next generation. In the selection process, we calculate each chromosome's probability which is in proportion to their fitness value. Then we calculate the cumulative probability. A number in [0, 1] is randomly generated. It will fall into a interval of the cumulative probability, and the corresponding chromosome is selected.

In the crossover process, we use single-point crossover operation. There are two parent chromosomes in Figure 2. We randomly generate a cross position colored bright yellow. Then they exchange part of their chromosomes from the cross position to generate child chromosomes. As shown in Figure 3, each child chromosome is formed by parts of two parent chromosomes. In the mutation process, we randomly generate a mutation position and flip the value(turn 0 into 1 or turn 1 into 0) on this position. The crossover operation and the mutation operation are employed for increasing the population diversity so that the search space is expanded. Both the crossover operation and the mutation operation are operated with probabilities. Traditional Genetic Algorithm usually set the crossover probability and the crossover probability as fixed values. However, recently they are always set as dynamic for better satisfying the demand of the algorithm in different period. In our research, the two probabilities are set as

$$P_c = \begin{cases} P_{cmax} - \frac{P_{cmax}-P_{cmin}}{k} \times k_{max}, & fit_l > fit_{ave} \\ P_{cmax}, & fit_l \leq fit_{ave} \end{cases} \tag{18}$$

$$P_m = \begin{cases} P_{mmax} - \frac{P_{mmax}-P_{mmin}}{k} \times k_{max}, & fit > fit_{ave} \\ P_{mmax}, & fit \leq fit_{ave} \end{cases} \tag{19}$$

where P_c and P_m are the crossover probability and the mutation probability, respectively; P_{cmax} and P_{cmin} are the maximum value and the minimum value of the crossover probability; P_{mmax} and P_{mmin} are the maximum value and the minimum value of the mutation probability; k is the current generation number; k_{max} is the maximum generation number; fit_l is the larger fitness value of two parent chromosome in the crossover operation; fit is the fitness value of the chromosome in the mutation operation; fit_{ave} is the average fitness value of current population.

The process of Genetic Algorithm used in this article is summarized below.

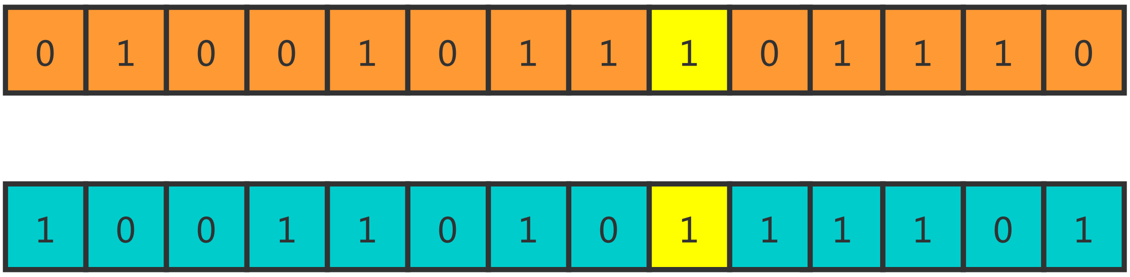

Figure 2: Parent chromosomes.

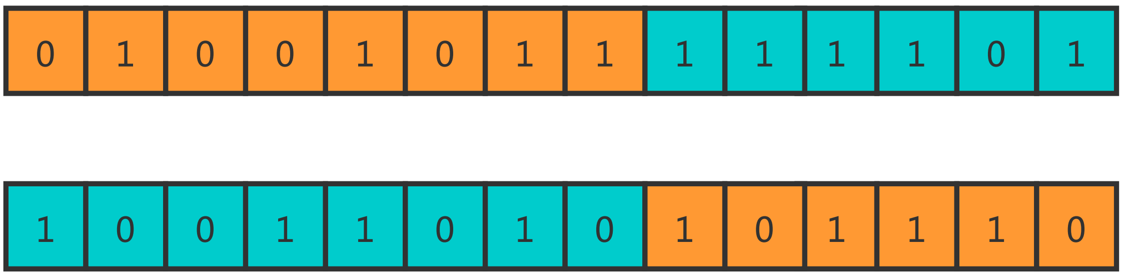

Figure 3: Child chromosomes.

Step 1: Initialize the population: determine the population size and the number of generations, randomly generate the initial populations X represented in binary codes.

Step 2: Calculate the fitness value of each individual $\Pi(X)$. Update and record the best fitness value and corresponding individual's binary codes.

Step 3: Select a pair of individuals from initial population. The individual with a higher fitness value could be chosen with a higher probability.

Step 4: Determine whether the crossover process will be conducted based on the crossover probability. The crossover process is to randomly choose a site of binary codes and to exchange parts of two selected individuals. Two new individuals are generated after the crossover process.

Step 5: Determine whether the mutation process will be conducted based on the mutation probability. The mutation process is to randomly choose a site of binary codes and to flip the bit in this site: 0 to 1, 1 to 0.

Step 6: Back to Step 3 until the new population of specified size is generated.

Step 7: Back to Step 2 until the number of generations is met.

4.3. Simulated Annealing

Simulated Annealing is a search method which is inspired by the physical annealing process. Kirkpatrick et al. (1983) employed the idea for solving optimization problems, where the objective function to be minimized corresponds to the internal energy of the metal. The difference between SA and the general local search methods is that SA accepts poor solutions with a small probability, rather than just selects the current best optimum solution into the next iteration, which would be helpful for avoiding local optimum. Starting from a higher temperature, the probability becomes lower with decreasing temperatures, and the solution gradually becomes stable. A simulated annealing algorithm has been presented to solve a location problem in Zarandi et al. (2013). We also develop a simulated annealing algorithm to solve the customer classification based location model in this paper. The process of Simulated Annealing used in this article is summarized below.

Step 1: Initialize the temperature T, the length of Markov chain L and initial solution $X_{CURRENT} = (x, \alpha)$.
Step 2: Generate new solution X_{NEW} according to the iterative formula.
Step 3: Calculate fitness values of the current solution $\Pi(X_{CURRENT})$ and of the new solution $\Pi(X_{NEW})$.
Step 4: Calculate the difference Δ between these two values. If $\Pi(X_{NEW})$ is higher than $\Pi(X_{CURRENT})$, accept X_{NEW} as $X_{CURRENT}$; otherwise accept X_{NEW} with a probability: $e^{-\Delta/T}$.
Step 5: For each temperature T, execute Step 2-4 L times.
Step 6: If the temperature T meets the the terminal condition, output $X_{CURRENT}$; otherwise reduce the temperature T and back to Step 2.

4.4. Tabu Search

The concept of Tabu Search was first proposed by Glover in 1986 and formulated in Glover (1989), and gradually developed into a mature algorithm. The main idea of Tabu Search is to record some solutions that have been found so as to avoid re-visiting these solutions and jump out the local optimum for the global optimum solution. Here, we also design a Tabu Search algorithm to solve the customer classification based location model. The process of Tabu Search used in this article is summarized below.

Step 1: Initialize the tabu list which is empty at first.
Step 2: Randomly generate a initial solution $X = (x, \alpha)$ and calculate its fitness value $\Pi(X)$. Assign X to the best solution X_{BEST} and the current solution $X_{CURRENT}$.
Step 3: Generate neighborhood solutions from neighborhood of the current solution. Calculate their fitness values. Choose the solution which has the largest fitness value as the candidate solution $X_{CANDIDATE}$.

Step 4: If the value of the candidate solution $\Pi(X_{CANDIDATE})$ is lower than of the current solution $\Pi(X_{CURRENT})$, assign $X_{CANDIDATE}$ to $X_{CURRENT}$ and update the tabu list. If $\Pi(X_{CANDIDATE})$ is higher than $\Pi(X_{CURRENT})$ and is higher than $\Pi(X_{BEST})$, assign $X_{CANDIDATE}$ to $X_{CURRENT}$ and X_{BEST}, and update the tabu list. If $\Pi(X_{CANDIDATE})$ is higher than $\Pi(X_{CURRENT})$ and is lower than $\Pi(X_{BEST})$, judge if $X_{CANDIDATE}$ is in the tabu list; if not, assign $X_{CANDIDATE}$ to $X_{CURRENT}$ and update the tabu list; otherwise apply $X_{CURRENT}$ regenerate neighborhood solutions.

Step 5: If the terminal condition is met, stop searching and output X_{BEST}, otherwise back to Step 3.

5. Computational results

In this section, we conduct numerical experiments on the customer classification based location model for comparing the performance of these four algorithms above in solving our model and evaluating the influence of customer proportion on the location choice.

5.1. Comparison of algorithms

As described above, we have designed four kinds of heuristic algorithms to solve the location model. The proposed algorithms were tested for a scenario where there are two competitive chains in the region, which may have certain numbers of existing facilities, respectively. Now one of the two chains is going to locate a new facility and determine the quality of the facility for maximizing the profit of the chain. We generate location problems with different settings, by varying the number of demand points n from 20 to 20000, and the number of existing facilities m from 5 to 300. For the convenience of computing, we randomly generate three possible customer proportions for each demand point. Then fuzzy estimated market share of each demand points is calculated according to (8). Following the parameter settings in Fernández et al. (2017a), $G_1(x)$ and $G_2(\alpha)$ in the location model are defined as

$$G_1(x) = \sum_{i=1}^{n} w_i \frac{1}{(d_i(x))^{\phi_{i0}} + \phi_{i1}},$$

and

$$G_2(\alpha) = e^{\frac{\alpha}{\alpha_0}+\alpha_1} - e^{\alpha_1}.$$

Parameters of the customer classification based location model are randomly chosen within the following intervals in Table 4.

The following ranges of PSO's parameters are tested according to the PSO literature:the particle number in $[5, 60]$, the acceleration coefficient in $[1, 4]$, and the inertia coefficient in $[0.4, 1.2]$ (Rezaee Jordehi and Jasni, 2013). We employ PSO with different parameter settings to solve the location problem in which the number of demand points is 5000, and the numbers

Table 4: Parameters in the location problems.

f_j	p_i	ω_i	α_{ij}	c	η	ϕ_{i0}	ϕ_{i1}	α_0	α_1
$[0, 10]^2$	$[0, 10]^2$	[1, 10]	[0.5, 5]	[1, 2]	[2, 2]	[2, 2]	[0.5, 2]	[7, 9]	[4, 4.5]

of existing facilities belonging to two competitive chains are 20 and 30, respectively. The optimal objective values obtained by PSO and the computing time (seconds) are shown in Table 5. In order to ensure the reliability of the results, the algorithm with each setting is operated 5 times. The maximum objective value and the average computing time of 5 times are recorded. The computing time is the time that the algorithm converges. The results shows that PSO with different settings obtains close objective values. We set the particle number, the acceleration coefficient and the inertia coefficient are 20, 2 and [0.4, 1] respectively, and employ this setting in our subsequent experiments.

Table 5: Parameter selection tests results of PSO.

Particle number	Acceleration coefficients	Inertia coefficient	Objective value	Computing time
10	1	[0.4, 1]	60663.07	4.34
10	1	[0.4, 1.2]	60662.91	6.164
10	2	[0.4, 1]	60663.1	8.552
20	2	[0.4, 1]	60663.11	15.918
20	2	[0.6, 1]	60663.11	22.11
20	1	[0.4, 1]	60663.07	9.7275
40	2	[0.4, 1]	60663.11	19.69

We also refer to Arostegui Jr et al. (2006) to test the performance of GA, SA and TS with different parameters. The results are shown in Table 6, 7, 8.

Table 6: Parameter selection tests results of GA.

Population number	Crossover probability	Mutation probability	Objective value	Computing time
10	[0.1, 0.9]	[0.01, 0.09]	49773.95	0.46
20	[0.1, 0.9]	[0.01, 0.09]	49908.87	1.61
20	[0.3, 0.9]	[0.03, 0.09]	49998.23	2.92
20	[0.5, 0.9]	[0.05, 0.09]	49466.14	5.20
40	[0.3, 0.9]	[0.03, 0.09]	50091.18	7.58
60	[0.3, 0.9]	[0.03, 0.09]	50160.35	15.32
80	[0.3, 0.9]	[0.03, 0.09]	50272.08	22.18
100	[0.3, 0.9]	[0.03, 0.09]	50101.61	52.23

Table 7: Parameter selection tests results of SA.

Initial temperature	Temperature decreasing rate	The length of markov chain	Objective value	Computing time
50	0.95	10	42918.58	11.14
70	0.95	10	43076.67	8.91
100	0.95	10	43623.24	12.06
100	0.98	10	43656.69	25.28
100	0.98	20	43670.34	60.03
150	0.98	20	43280.49	63.42

Table 8: Parameter selection tests results of Tabu.

Minimum Tabu list size	Maximum Tabu list size	Objective value	Computing time
5	11	43271.52	15.48
5	15	43291.98	15.66
5	20	43292.93	14.98
10	20	43299.58	11.53
10	30	43278.82	6.37
15	20	60231.77	14.69

Based on experimental results, the parameters of these algorithms are set as Table 9. Next we employ these four algorithms with parameters listed in Table 9 to solve the location problems. We run these four algorithms in the personal computer environment: Intel i3-3420 CPU, 3.40 GHz, 4GB RAM and Windows 10 system.

To ensure the reliability of these four heuristic search algorithms, each algorithm for each problem is operated 10 times. The average objective values and average computing time (seconds) are recorded in Table 10. The first column lists the setting (n, j^k_{max}) of each problem. For example, The problem with setting of (50, (6,10)) means that in this problem the number of demand pints is 50, and the numbers of existing facilities belonging to two competing chains are 6, 4, respectively. The chain with 3 existing facilities is going to locate a new facility. For each algorithm, two columns are listed recording the average objective values and average computing time when solving corresponding problems, respectively. From the results we get the following conclusions:

(1) For small size problems, the objective values employing four algorithms are almost equal.
(2) When the problem size becomes larger, the gaps between the objective values obtained by GA and by other three algorithms are getting more significant.
(3) The objective value obtained by SA is closer to the value obtained by PSO, but the computing time spent by employing SA is far longer than by employing PSO.
(4) For all the problems we have generated, PSO is the algorithm that always provides the results of the best quality in the shortest time among these four algorithms.

Table 9: Parameters in heuristic algorithms.

PSO	Particle number	Acceleration coefficients	Inertia coefficient
	20	2	[0.4, 1]
GA	Population number	Crossover probability	Mutation probability
	80	[0.3, 0.9]	[0.03, 0.09]
SA	Initial temperature	Temperature decreasing rate	The length of markov chain
	100	0.98	20
TS	Tabu list size		
	[15, 20]		

Table 10: The computation results of location problems employing different algorithms.

Settings	PSO		GA		SA		TS	
	Objective	Time (s)	Objective	Time (s)	Objective	Time (s)	Objective	Time (s)
(20,(3,5))	453.61	0.18	450.58	0.18	453.60	4.12	453.24	0.90
(50,(6,10))	860.55	0.12	859.90	0.36	860.55	7.02	860.15	1.04
(100,(10,20))	1903.33	0.17	1858.51	0.31	1903.31	11.65	1902.90	1.80
(1000,(20,50))	11981.76	1.39	11969.17	4.98	11981.08	90.69	11980.08	17.20
(3000,(40,100))	34921.86	10.60	34831.54	25.97	34918.39	698.60	34910.48	112.48

After observing the performance of these four algorithms in solving a set of location problems we generated, we select the problem with setting of (3000, (40,100)) to analyze the stability of these algorithms. The obtained maximum value, minimum value, average value as well as the standard deviation of the objective value and the computing time of the location problem with the setting of (3000, (40,100)) are recorded in Table 11. It clearly shows PSO is superior to other three heuristics for the location problem we studied both from the computing time and the solution quality two aspects. The standard deviation of objective value obtained by PSO is 0.25, which indicates the robustness of PSO. The optimal objective value obtained by SA is near to the value of PSO, however it's standard deviation is 513.15, which means SA's performance is unstable. The objective value obtained by GA and TS also have different degrees of fluctuation.

Table 11: Detailed computation results of a problem using different algorithms.

	PSO		GA		SA		TS	
	Objective	Time (s)	Objective	Time (s)	Objective	Time (s)	Objective	Time (s)
Max	34921.86	13.21	34831.54	54.56	34918.39	704.73	34910.48	193.18
Min	34921.28	7.13	34433.56	3.57	33661.48	695.29	34855.88	39.97
Average	34921.72	10.60	34676.29	25.97	34286.49	698.60	34891.92	112.48
Standard deviation	0.25	2.22	138.24	16.62	513.15	4.34	18.75	53.98

Overall, PSO performs best in solving the location model we proposed. In addition, PSO is able to handle large size problems. It is the most suitable algorithm to our problem comprehensively considering the solution quality, algorithm stability and computation time. In the following tests, we adopt PSO as the acceptable solution approach to solve the customer classification based location model.

5.2. The influence of customer proportion

In this subsection, we study the influence of customer proportion on the location of the new facility and on the profit obtained by the locating chain. We conduct the experiments with fixed customer proportions and random customer proportions respectively to study the influence of customer proportion.

For comparing the location results obtained by employing different customer proportions, we consider following location strategies:

(1) Deterministic strategy: All the customers patronize facilities with the deterministic rule. Based on this assumption locating chain makes the location decision.
(2) Probabilistic strategy: All the customers patronize facilities with the probabilistic rule. Based on this assumption locating chain makes the location decision.
(3) Multi-deterministic strategy: All the customers patronize facilities with the multi-deterministic rule. Based on this assumption locating chain makes the location decision.
(4) Comprehensive strategy: Different customers patronize facilities with different choice rules. For each demand point, the market share is calculated based on the customer proportion. The market share is estimated by real variables. Based on this assumption locating chain makes the location decision.
(5) Fuzzy comprehensive strategy: Different customers patronize facilities with different choice rules. For each demand point, the fuzzy market share is calculated based on several possible customer proportions. The market share is estimated by fuzzy variables. Based on this assumption locating chain makes the location decision.

We denote Π_D as the objective function of the location problem when the deterministic strategy is employed, denote x^*_D and α^*_D as the optimal location coordinates and quality of the new facility found by PSO algorithm, denote Π_P, x^*_P and α^*_P as the corresponding items when the probabilistic strategy is employed, denote Π_M, x^*_M, and α^*_M as the corresponding items when the multi-deterministic strategy is employed, denote Π_C, x^*_C, and α^*_C as the corresponding items when the comprehensive strategy is employed, and denote Π_F, x^*_F, and α^*_F as the corresponding items when the fuzzy comprehensive strategy is employed.

In order to compare the location results obtained by using different customer proportions, we compute the Euclidean distance between different location results and denote it by $disl$. For example, $disl_{DP}$ represents the distance between x^*_D and x^*_P. Similarly, the difference between the quality is computed and denoted by $disq$. Besides, the relative profit loss of locating chain also needs to be computed when a wrong customer choice rule proportion is applied. For instance, when we employ the optimal location results of assumption (4) in the assumption (5), we compute the relative loss

$$loss(M|F) = 100 * (\Pi_F(x^*_F, \alpha^*_F) - \Pi_F(x^*_M, \alpha^*_M))/\Pi_F(x^*_F, \alpha^*_F)$$

to measure the loss of the locating chain caused by employing the wrong customer proportion.

We randomly generate 50 demand points and 10 existing facilities in the market area which is a 10×10 square. For each demand point, 3 possible customer proportions are generated. Fuzzy market share of each demand point is calculated based on customer proportions.

For observing the results under various market situations, we assume several representative competitive market conditions.

Scenario 1 "newcomer 1": $k = 2$ (number of chains), and the number of existing facilities belonging to each chain is 0 and 10, respectively, which means that the locating chain (chain 1) is a newcomer to the market and all the existing facilities belong to another competing chain.

Scenario 2 "newcomer 2": $k = 3$, and the number of existing facilities belonging to each chain is 0, 4 and 6, respectively. Similarly, the locating facility has no existing facility. But different with scenario "newcomer 1", existing facilities belong to two chains in this case. Competition exists before the new chain enter the market.

Scenario 3 "small chain": $k = 2$, and the number of existing facilities belonging to each chain is 3 and 7, respectively. The locating chain has 3 existing facilities.

Scenario 4 "large chain": $k = 2$, and the number of existing facilities belonging to each chain is 7 and 3, respectively. The locating chain has 7 existing facilities.

The location problems are solved by PSO algorithm introduced in Section 4.1. For each test, we observe the convergence graph to make sure the algorithm is converged so the optimal solution can be obtained. We find most tests converged in 100 generations, so we set the default number of iteration generations as 100, and adjust the parameter depending on the specific tests.

The results obtained are given in Tables 12, 13, and 14. From the Tables 12 and 13 we can see that the location results obtained by using different customer proportions show differences in various degree. In our location problem, as a result of the limitation of decision variables, the maximum value of the distance between different locations is $\sqrt{200}$ (approximately equal to 14.15), the maximum value of the difference between different quality level is 4.5. In our numerical experiment, the maximum values of the two values are 5.88 and 1.24, respectively. As we can see, the customer proportion has a significant influence in location results. Different customer proportions can lead to quite different location results. Notice that when the market share is estimated by fuzzy variables, the location results also have an obvious difference compared with other location results.

Table 14 displays the relative profit loss caused by employing wrong customer proportions in the assumption (5). As we can see, the relative profit loss incurred for the chain when a wrong customer proportion is employed can be quite large in some instances. The relative profit loss is 12.76% when the locating chain wrongly assumes all the customers are the multi-deterministic type in the Scenario 1 "newcomer 1", which is a rather high profit loss for the chain. In addition, we found that the relative profit loss in the Scenario 1 "newcomer 1" and Scenario 2 "newcomer 2" is higher than in the latter two scenarios.

As we have seen in this case study, the location results can be quite different by employing different customer assumptions. The chain may lose large profit when a wrong customer

Table 12: Differences of the new facility's location caused by different customer assumptions.

Scenario	$distl_{DP}$	$distl_{DM}$	$distl_{DC}$	$distl_{DF}$	$distl_{PM}$
1	2.48	2.52	0.13	3.58	4.91
2	2.48	2.06	0.13	3.58	0.42
3	2.30	3.09	1.21	3.11	2.97
4	0.46	0.50	0.47	0.47	0.04
Scenario	$distl_{PC}$	$distl_{PF}$	$distl_{MC}$	$distl_{MF}$	$distl_{CF}$
1	2.52	5.88	2.44	1.12	3.49
2	2.52	5.88	2.11	5.50	3.49
3	1.37	2.99	3.40	0.02	3.42
4	0.01	0.01	0.03	0.03	0.00

Table 13: Differences of the new facility's quality caused by different customer assumptions.

Scenario	$distq_{DP}$	$distq_{DM}$	$distq_{DC}$	$distq_{DF}$	$distq_{PM}$
1	0.09	0.09	0.09	0.09	0.00
2	0.09	0.09	0.09	0.09	0.00
3	1.02	1.02	1.02	1.02	0.00
4	1.24	1.24	1.24	1.24	0.00
Scenario	$distq_{PC}$	$distq_{PF}$	$distq_{MC}$	$distq_{MF}$	$distq_{CF}$
1	0.00	0.00	0.00	0.00	0.00
2	0.00	0.00	0.00	0.00	0.00
3	0.00	0.00	0.00	0.00	0.00
4	0.00	0.00	0.00	0.00	0.00

assumption is used. Considering that different customers may act with different customer choice rules, a chain should investigate the proportion of customers using several kinds of choice rules before locating a new chain in order to avoid high profit loss, especially for the newcomer of the market.

6. Conclusions and future research

The main contribution of this paper is to develop a customer classification based competitive facility location model to help finding a better location for the new facility and improving the profit of the locating chain. First, customers were classified into three types: the deterministic type, the probabilistic type and the multi-deterministic type. Then a customer classification based facility location model in the plane was proposed in which location and quality of the new facility are to be determined in order to maximize the profit of the locating chain. Furthermore, we designed PSO algorithm to solve this model and

Table 14: The relative profit loss caused by employing wrong customer assumptions in the assumption (5).

Scenario	$loss(D\|F)$	$loss(P\|F)$	$loss(M\|F)$	$loss(C\|F)$
1	4.22	8.99	12.76	3.28
2	5.40	6.64	6.40	4.38
3	2.46	3.56	0.01	0.02
4	1.65	0.00	0.00	0.00

compared the performance of PSO with GA, SA, and TS. By comparison and analysis, PSO showed its superior performance in solving our problem. A set of computational tests were conducted to study the influence of customer proportion on location choice and to evaluate profit loss of the locating chain when a wrong customer proportion is employed.

According to the computational results, the optimal location and optimal quality of the new facility and the profit obtained by the locating chain varied greatly when different customer proportions were employed. In our computational tests, the locating chain suffered the profit loss of different extent when a wrong customer proportion was employed. The maximum relative profit loss was 12.76%, which was a great loss to the chain. In many cases the relative profit loss was over 5%. The employment of a wrong customer proportion could bring a huge profit loss to the locating chain, especially for a newcomer. Hence the customer proportion should be investigated before a facility is located.

Our work can be extended to the location problem on the network. The location problem in the discrete location space can be formulated as an integer programming model and the location results may also be different. The extension of our work to the case of the location of multi-facilities also need to be discussed, which will add the computing complexity with the increase of variables. Improving the precision of the algorithm that we used to solve the location problem is also a main direction in our future research.

Acknowledgement

This work was supported by grants from the National Natural Science Foundation of China (Nos. 71722007 & 71931001), the Funds for First-class Discipline Construction (XK1802-5) and the Fundamental Research Funds for the Central Universities (buctrc201926).

References

Arostegui Jr, M. A., Kadipasaoglu, S. N., Khumawala, B. M., 2006. An empirical comparison of tabu search, simulated annealing, and genetic algorithms for facilities location problems. International Journal of Production Economics 103 (2), 742–754.

Biesinger, B., Hu, B., Raidl, G., 2016. Models and algorithms for competitive facility location problems with different customer behavior. Annals of Mathematics and Artificial Intelligence 76 (1-2), 93–119.

Blanquero, R., Carrizosa, E., Boglárka, G., Nogales-Gómez, A., 2016. p-facility huff location problem on networks. European Journal of Operational Research 255 (1), 34–42.

Blanquero, R., Carrizosa, E., Nogales-Gómez, A., Plastria, F., 2014. Single-facility huff location problems on networks. Annals of Operations Research 222 (1), 175–195.

da Gama, F. S., Captivo, M. E., 1998. A heuristic approach for the discrete dynamic location problem. Location science 6 (1-4), 211–223.

Dan, T., Marcotte, P., 2019. Competitive facility location with selfish users and queues. Operations Research 67 (2), 479–497.

Drezner, T., 1994a. Locating a single new facility among existing, unequally attractive facilities. Journal of Regional Science 34 (2), 237–252.

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

Drezner, T., 1998. Location of multiple retail facilities with limited budget constraintsin continuous space. Journal of Retailing and Consumer Services 5 (3), 173–184.

Drezner, T., 2009. Location of retail facilities under conditions of uncertainty. Annals of Operations Research 167 (1), 107–120.

Drezner, T., 2014. A review of competitive facility location in the plane. Logistics Research 7 (1), 114.

Drezner, T., Drezner, Z., 2004. Finding the optimal solution to the huff based competitive location model. Computational Management Science 1 (2), 193–208.

Drezner, T., Drezner, Z., Kalczynski, P., 2011. A cover-based competitive location model. Journal of the Operational Research Society 62 (1), 100–113.

Eberhart, R., Kennedy, J., 1995. Particle swarm optimization. In: Proceedings of the IEEE international conference on neural networks. Vol. 4. Citeseer, pp. 1942–1948.

Eiselt, H. A., Laporte, G., 1997. Sequential location problems. European Journal of Operational Research 96 (2), 217–231.

Fernández, J., Boglárka, G., Redondo, J. L., Ortigosa, P. M., 2019. The probabilistic customers choice rule with a threshold attraction value: Effect on the location of competitive facilities in the plane. Computers & Operations Research 101, 234–249.

Fernández, J., Boglárka, G., Redondo, J. L., Ortigosa, P. M., Arrondo, A. G., 2017a. A planar single-facility competitive location and design problem under the multi-deterministic choice rule. Computers & Operations Research 78, 305–315.

Fernández, J., Pelegrín, B., Plastria, F., Tóth, B., 2007. Solving a huff-like competitive location and design model for profit maximization in the plane. European Journal of Operational Research 179 (3), 1274–1287.

Fernández, P., Pelegrín, B., Lančinskas, A., Žilinskas, J., 2017b. New heuristic algorithms for discrete competitive location problems with binary and partially binary customer behavior. Computers & Operations Research 79, 12–18.

Glover, F., 1989. Tabu searchpart i. ORSA Journal on computing 1 (3), 190–206.

Hakimi, S. L., 1983. On locating new facilities in a competitive environment. European Journal of Operational Research 12 (1), 29–35.

Hakimi, S. L., 1990. Locations with spatial interactions: competitive locations and games. Discrete location theory.

Holland, J. H., 1992. Adaptation in Natural and Artificial Systems. MIT Press, Cambridge, MA, USA.

Hotelling, H., 1990. Stability in competition. In: The Collected Economics Articles of Harold Hotelling.

Springer, pp. 50–63.

Huff, David, L., 1964. Defining and estimating a trade area. Journal of Marketing 28, 34–38.

Kirkpatrick, S., Gelatt, C. D., Vecchi, M. P., 1983. Optimization by simulated annealing. science 220 (4598), 671–680.

Lančinskas, A., Fernández, P., Pelegín, B., Žilinskas, J., 2017. Improving solution of discrete competitive facility location problems. Optimization Letters 11 (2), 259–270.

Liu, B., Liu, Y.-K., 2002. Expected value of fuzzy variable and fuzzy expected value models. IEEE transactions on Fuzzy Systems 10 (4), 445–450.

Nakanishi, M., Cooper, L. G., 1974. Parameter estimation for a multiplicative competitive interaction modelleast squares approach. Journal of marketing research 11 (3), 303–311.

Özgün-Kibiroğlu, Ç., Serarslan, M. N., Topcu, Y. İ., 2019. Particle swarm optimization for uncapacitated multiple allocation hub location problem under congestion. Expert Systems with Applications 119, 1–19.

Plastria, F., 2001. Static competitive facility location: an overview of optimisation approaches. European Journal of Operational Research 129 (3), 461–470.

Redondo, J. L., Fernández, J., Arrondo, A. G., García, I., Ortigosa, P. M., 2012. Fixed or variable demand? does it matter when locating a facility? Omega 40 (1), 9–20.

Redondo, J. L., Fernández, J., García, I., Ortigosa, P. M., 2009. A robust and efficient algorithm for planar competitive location problems. Annals of Operations Research 167 (1), 87–105.

Rezaee Jordehi, A., Jasni, J., 2013. Parameter selection in particle swarm optimisation: a survey. Journal of Experimental & Theoretical Artificial Intelligence 25 (4), 527–542.

Serra, D., Colomé, R., 2001. Consumer choice and optimal locations models: formulations and heuristics. Papers in Regional Science 80 (4), 439–464.

Shan, W., Yan, Q., Chen, C., Zhang, M., Yao, B., Fu, X., 2019. Optimization of competitive facility location for chain stores. Annals of Operations Research 273 (1-2), 187–205.

Suárez-Vega, R., Santos-Peñate, D. R., Dorta-González, P., 2004. Competitive multifacility location on networks: the (ro xp)-medianoid problem. Journal of Regional Science 44 (3), 569–588.

Zarandi, M. H. F., Davari, S., Sisakht, S. A. H., 2013. The large-scale dynamic maximal covering location problem. Mathematical and Computer Modelling 57 (3-4), 710–719.