An e-commerce facility location problem under uncertainty

J. Shen*

Department of Public Basic Courses, Nanjing Institute of Industry Technology, Nanjing 210023, Jiangsu, People’s Republic of China

Received 11 February 2018; received in revised form 28 March 2019; accepted 20 May 2019

KEYWORDS
Facility location; Expected value; Chance-constrained; Uncertain environment; Supply chain network.

Abstract. Facility location problem is a branch of operations research and computational geometry. It covers the best allocation of facilities to minimize transportation costs by considering the factors involved (e.g., avoiding the placement of dangerous materials near the premises and the facilities of competitors). Given the unique customer characteristics and the fierce market competition of business-to-consumer e-commerce, the expected value model and chance-constrained model for uncertain facility location problems were constructed. Owing to the intricacies of the competitive market, supply capacity, delivery cost, and customer demand were assumed as uncertain variables. The deterministic equivalent forms of the models were discussed using the inverse uncertainty distribution method. A hybrid algorithm was proposed to solve these models. Some numerical experiments were used to verify the effectiveness of the proposed models and method.

© 2021 Sharif University of Technology. All rights reserved.

1. Introduction

Business-to-Consumer (B2C), as one of the three modes of e-commerce, has become the core force to promote the online shopping market. The market size of B2C has exceeded that of Consumer-to-Consumer (C2C), which was expected to reach 70% by 2018. However, unlike the conventional commercial distribution characterized by high volume and small batches, B2C is no longer a retailer but a direct response to numerous customers with low demand, rich variety and scattered locations. Distribution business has many characteristics including many customers, wide distribution, many varieties, and small batch, thereby resulting in a complex logistics system, high cost, low service levels, and other issues. For online shopping, customers can only judge the product quality based on the basic information of products, customer evaluation, etc. When the quality assurance is not high, coupled with the logistics damage, loss, etc., many goods will be returned. B2C logistics system presents a totally uncoordinated development trend with the fast-growing online shopping market, thereby limiting the further development of enterprises and arousing the attention of many scholars [1–3].

In recent years, many studies have been conducted on the location problem. Klose and Drex [4] reviewed some facility location models and solution algorithms for distribution system design. Manzini and Gebennini [5] developed an innovative model for location assignment problem in a distribution system. Lau et al. [6] proposed a fuzzy B2C location model and an improved hybrid algorithm was used to solve this model. Chen et al. [7] proposed a location-inventory model with facility destruction and a Lagrangian relaxation solution framework. Berman et al. [8] proposed a location inventory model and a Lagrangian relaxation algorithm. Tancrez et al. [9] analyzed a three-level location-inventory problem. It was proved that when the Distribution Center (DC) flow was fixed, it could

* E-mail address: jcyue007@126.com (J. Shen)

be decomposed into a closed equation and a linear programming. Shahabi et al. [10] considered a three-level location-inventory problem with the correlated demand. A novel scheme to convert the initial formulation into mixed integer conic programming and an outer approximation strategy were proposed. The research on distribution systems has not been enough as it can significantly affect the profitability of the range-saving companies by 5% to 10%. Rashidi et al. [11] studied a perishable-item location-inventory problem. A bi-objective mathematical model was developed and a Pareto-based meta-heuristic method was employed to solve the model. Lin et al. [12] studied a multi-classification-yard location problem and used an efficient simulated annealing algorithm to solve the problem. Labbé et al. [13] considered a hierarchical location problem with two types of facilities and developed alternative Benders decomposition algorithms.

B2C e-commerce allows for the direct trade between enterprises and customers. To save the cost of distribution, enterprises should operate their own fleet on a line to serve services to many customers. Under these circumstances, the cost of delivery is hard to estimate accurately. In recent years, because of the unpredictability of logistics and distribution system, decision-makers have been facing uncertain events frequently. Accordingly, the facility location problem in a random environment has aroused huge attention. Snyder et al. [14] studied a stochastic location model with risk pooling, which is used to minimize the expected value of the total cost. Tezenji et al. [15] developed an integrated model for a facility location-allocation problem. Genetic Algorithm (GA) and simulated annealing were employed to solve the mixed-integer nonlinear program. Marković et al. [16] proposed the first multi-period stochastic flow-capturing model for facility location problem and a Lagrangian relaxation algorithm. Amiri-Aref et al. [17] proposed a two-stage stochastic mathematical model for the location-inventory problem. A linear approximation was employed to obtain near-optimal solutions.

The fuzzy theory provided by Zadeh [18] can be an alternative method to address the facility location problem. Some research studies were conducted in the area of facility location modeling with fuzzy parameters [19–23].

It is worth mentioning that the methods mentioned in the above literature cannot be directly used to solve the problem of the uncertain B2C e-commerce facility location. First, the characteristics of the customer have not been considered, e.g., small batch demand and geographically dispersed locations. Second, the above research pieces in the literature mostly have focused on how to configure the location and quantity of DCs, ignoring their capacity and always assuming that the capacity is fixed.

It is generally known that the precondition of using probability theory is that the probability distribution is available. In a random environment, the random variables can be estimated based on the historical data. However, in many facility location problems, probability distributions are often not available due to the lack of the accurate data. In this case, experts can only assess the degree of belief that whether uncertain events will occur. The degree of belief is largely determined by a large extent on personal experience. To deal with the degree of belief, uncertainty theory was initiated in [24] and refined in [25]. Uncertainty theory is a useful tool for solving such problems in an uncertain environment. Uncertainty theory is a branch of axiomatic mathematics for modeling human uncertainty, which has many research results, e.g., uncertain programming [26–29], uncertain risk analysis [30–32], uncertain calculus [33–35], and uncertain differential equation [36–38].

The problem of an uncertain facility location in B2C e-commerce was studied here. In reality, some factors (e.g., demands and locations of customers, allocations, and facilities) are usually changing. To make a better decision, decision-makers may consider more complex situations. Thus, it is of great practical implication to study the uncertain facility location problem. The aim is to minimize the total logistics cost under an uncertain environment. Moreover, for small-scale problems, the expected value model and chance-constrained model are developed. It is proved that the models can be converted into crisp models. Finally, an efficient hybrid intelligent algorithm integrating GA and Particle Swarm Optimization (PSO) is proposed based on the theoretical analysis and the characteristics of the deterministic models.

The rest of the study is organized as follows. Section 2 briefly introduces the uncertainty theory. Section 3 describes the concern in this study and constructs two models in the uncertain environment. Section 4 discusses the equivalence of models. Section 5 proposes a hybrid intelligent algorithm. Section 6 performs numerical experiments to illustrate the validity of the proposed models and algorithm.

2. Preliminaries

A brief introduction to uncertainty theory is given. To describe an uncertain variable which refers to human uncertainty, Liu [24] established the uncertainty theory, which has been developed well up to now.

Let Ω be a nonempty set, ℳ be a σ-algebra over Ω, and each element Λ in Λ be called an event. A set function μ from ℳ to [0, 1] is called an uncertain measure if it satisfies normality axiom, duality axiom, subadditivity axiom, and product axiom [24,39].

An uncertain variable is a measurable function ξ
from an uncertainty space \((\Gamma, \mathcal{L}, \mathcal{M})\) to the set \(R\) of real numbers, i.e., for any Borel set \(B\) of real numbers and the set \(\{\xi \in B\} = \{\gamma \in \Gamma | \xi(\gamma) \in B\}\) is an event. The distribution \(\Phi\) of an uncertain variable \(\xi\) is defined by \(\Phi(x) = \mathcal{M}\{\xi \leq x\}\) for any real number \(x\). The uncertain variables \(\xi_1, \xi_2, \cdots, \xi_m\) are said to be independence [39] if:

\[
\mathcal{M}\left\{\bigcap_{i=1}^{m} B_i \right\} = \min_{1 \leq i \leq m} \mathcal{M}\{\xi_i \in B_i\},
\]

for any Borel sets \(B_1, B_2, \cdots, B_m\) of real numbers.

**Definition 1** [24]. Let \(\xi\) be an uncertain variable and \(\alpha \in (0, 1]\). Then,

\[
\xi_{\text{sup}}(\alpha) = \sup\{r | \mathcal{M}\{\xi \geq r\} \geq \alpha\}
\]

is called the \(\alpha\)-optimistic value to \(\xi\) and:

\[
\xi_{\text{inf}}(\alpha) = \inf\{r | \mathcal{M}\{\xi \leq r\} \geq \alpha\}
\]

is called the \(\alpha\)-pessimistic value to \(\xi\).

**Definition 2** [24]. An uncertain distribution \(\Phi(x)\) is said to be regular if its inverse function \(\Phi^{-1}(x)\) exists and is unique for each \(\alpha \in (0, 1]\). Then, the inverse function \(\Phi^{-1}\) is called the inverse uncertainty distribution of \(\xi\).

**Example 1.** Let \(\xi_1, \xi_2, \cdots, \xi_n\) be independent and positive uncertain variables with regular uncertainty distributions \(\Phi_1, \Phi_2, \cdots, \Phi_n\), respectively. It can be shown that the product:

\[
\xi = \xi_1 \times \xi_2 \times \cdots \times \xi_n
\]

has an inverse uncertainty distribution as follows:

\[
\Phi^{-1}(\alpha) = \Phi_1^{-1}(\alpha) \times \Phi_2^{-1}(\alpha) \times \cdots \times \Phi_n^{-1}(\alpha).
\]

**Theorem 1** [25]. Assume that \(\xi_1, \xi_2, \cdots, \xi_n\) are independent uncertain variables with regular uncertainty distributions \(\Phi_1, \Phi_2, \cdots, \Phi_n\), respectively. If \(f(x_1, x_2, \cdots, x_n)\) is strictly increasing with respect to \(x_1, x_2, \cdots, x_m\) and strictly decreasing with respect to \(x_{m+1}, x_{m+2}, \cdots, x_n\), then the uncertain variable \(\xi = f(\xi_1, \xi_2, \cdots, \xi_n)\) has an expected value:

\[
E[\xi] = \int_{\alpha}^{1} f(\Phi_1^{-1}(\alpha), \cdots, \Phi_m^{-1}(\alpha), \Phi_{m+1}^{-1}(1-\alpha), \cdots, \Phi_n^{-1}(1-\alpha)) d\alpha
\]

provided that \(E[\xi]\) exists.

For any real numbers \(a\) and \(b\), we have \(E[a \xi + b \eta] = aE[\xi] + bE[\eta]\), where \(\xi\) and \(\eta\) are independent of each other.

### 3. Uncertain facility location models

#### 3.1. Description

A distribution system is considered here. There are a B2C company, several vendors, several potential DCs, and several customers in the system. The B2C company orders goods from suppliers, and suppliers deliver goods to the DCs directly. Different DCs can cope with different customer zones. To optimize the entire system, two models of uncertain facility location problems are established. The objective is to select the optimal quantity, location, and capacity of DCs so that the total cost can be minimized while meeting the demands of customers. The total cost covers the supply cost, transportation cost from the supplier to the DC, installation cost of the DC, inventory cost and management cost of the DC, and cost of delivery from the DC to the customer. There are some assumptions as follows:

- The distribution system includes a group of suppliers and customers with known locations as well as potential locations for DCs. Each customer can only obtain the goods from one DC;
- The demand of a customer, the capacity of a supplier, and the delivery cost are considered uncertain variables. The locations of suppliers and customers are fixed;
- The planning period includes several transport cycles and the goods are shipped from the supplier to the DC. Besides, each shipping cycle includes several similar delivery cycles in which the items are shipped from the DC to the customer.

#### 3.2. Mathematical models under uncertainty

Before building the mathematical models, parameters and variables are given as follows:

**Indexes and parameters**

<table>
<thead>
<tr>
<th>Index</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i)</td>
<td>Index of supplier, (i = 1, 2, \cdots, I)</td>
</tr>
<tr>
<td>(j)</td>
<td>Index of potential DC, (j = 1, 2, \cdots, J)</td>
</tr>
<tr>
<td>(k)</td>
<td>Index of customer, (k = 1, 2, \cdots, K)</td>
</tr>
<tr>
<td>(l)</td>
<td>Index of commodity, (l = 1, 2, \cdots, L)</td>
</tr>
<tr>
<td>(m)</td>
<td>Index of transport period, (m = 1, 2, \cdots, M)</td>
</tr>
<tr>
<td>(n)</td>
<td>Index of distribution period, (n = 1, 2, \cdots, N)</td>
</tr>
<tr>
<td>(A_i)</td>
<td>Supply capability of supplier (i) for the commodity (l)</td>
</tr>
<tr>
<td>(B_{jk})</td>
<td>Unit delivery cost from DC (j) to customer (k)</td>
</tr>
<tr>
<td>(C_{ik})</td>
<td>Unit shipping cost of the commodity (l) from supplier (i) to DC (j)</td>
</tr>
<tr>
<td>(D_{kl})</td>
<td>Demand of customer (k) for the commodity (l)</td>
</tr>
</tbody>
</table>
under the criterion of expected value. The model is as follows:

\[
\begin{align*}
\min E \left\{ m \sum_{i \in I} \sum_{l \in L} H_l \sum_{j \in J} x_{ij} + m \sum_{i \in I} \sum_{j \in J} \sum_{l \in L} C_{ijl} x_{ijl} \\
+ \sum_{j \in J} F_j \left( \sum_{i \in I} \sum_{l \in L} w_{ij} x_{ij} \right) \right. \\
= m \sum_{i \in I} \sum_{j \in J} \sum_{l \in L} n \phi \\
U_j \frac{x_{ij}}{n} + m \sum_{j \in J} \sum_{l \in L} \left( \sum_{i \in I} x_{ij} \right) S_{ij} \\
+ \sum_{j \in J} B_{jk} z_{jk} \left( \sum_{l \in L} q_l D_{lk} \right) \left. \right\},
\end{align*}
\]

subject to:

\[
E \left[ m \sum_{j \in J} x_{ijl} - A_{il} \right] \leq 0, \quad i \in I, \quad l \in L, \quad (2)
\]

\[
E \left[ m \sum_{i \in I} x_{ijl} - \sum_{k \in K} D_{lk} z_{jk} \right] = 0, \quad j \in J, \quad l \in L, \quad (3)
\]

\[
\sum_{i \in I} \sum_{l \in L} w_{ij} x_{ij} - M_j y_j \leq 0, \quad j \in J, \quad (4)
\]

\[
\sum_{j \in J} y_j - P \leq 0, \quad (5)
\]

\[
\sum_{j \in J} z_{jk} = 1, \quad k \in K, \quad (6)
\]

\[
\sum_{k} z_{jk} - r y_j \leq 0, \quad j \in J, \quad (7)
\]

\[
x_{ijl} \geq 0, \quad y_j = \{0, 1\}, \quad z_{jk} = \{0, 1\}, \quad i \in I, \quad j \in J, \quad l \in L. \quad (8)
\]

Inequality (2) ensures that the goods supplied by the supplier do not exceed their ability. Eq. (3) ensures that the input of each DC is equal to the output. Inequality (4) indicates that each DC cannot exceed its maximum capacity limit. Inequality (5) ensures that the selected DCS would not exceed the maximum number. Eq. (6) ensures that each customer can only obtain the goods from one DC. Inequality (7) ensures that each DC can deliver goods to customers.

In practice, the decision-maker always considers the risk and finds an upper bound to make an optimal schedule plan. Under different conditions, confidence levels \( \alpha \) are given. The decision-maker should determine target \( \alpha \) such that a solution \( x^* \) could satisfy...
\[ M\{f(x) \leq \bar{f}\} \geq \alpha. \] For instance, set \( \alpha = 0.9 \), the decision-maker should determine a target \( \bar{f} \) and then, choose a solution \( x \) that satisfies \( M\{f(x) \leq \bar{f}\} \geq 0.9 \). This suggests that if the decision-maker chooses \( x \), the total cost will be lower than \( \bar{f} \) at least 90%.

Accordingly, a chance-constrained model is conceived.

\[
\min \bar{f} \tag{9}
\]

subject to:

\[
M\left\{ m \sum_{i \in I} \sum_{l \in L} H_{jl} \sum_{i \in I} x_{ijl} + m \sum_{i \in I} \sum_{j \in J} \sum_{l \in L} C_{ijl} x_{ijl} + \sum_{j \in J} F_{j}(\sum_{i \in I} \sum_{l \in L} \sum_{j \in J} w_{ijl}) + m \sum_{i \in I} \sum_{j \in J} \sum_{l \in L} r_{i}^{j} S_{jl} \right\} \leq \bar{f} \tag{10}
\]

subject to:

\[
M\left\{ m \sum_{i \in I} x_{ijl} \leq A_{il}, \quad l \in L, \right\} \tag{11}
\]

\[
M\left\{ m \sum_{i \in I} x_{ijl} = D_{kl} z_{jk}, \quad l \in L, \right\} \tag{12}
\]

\[
\sum_{i \in I} \sum_{l \in L} w_{ijl} - M_{jl} y_{j} \leq 0, \quad j \in J, \tag{13}
\]

\[
\sum_{j \in J} y_{j} - P \leq 0, \tag{14}
\]

\[
\sum_{j \in J} z_{jk} = 1, \quad k \in K, \tag{15}
\]

\[
\sum_{k} z_{jk} - r y_{j} \leq 0, \quad j \in J, \tag{16}
\]

\[
x_{ijl} \geq 0, \quad y_{j} = \{0, 1\}, \quad z_{jk} = \{0, 1\}, \tag{17}
\]

where \( \alpha, \beta_1, \beta_2 \) are the preset confidence levels.

The model aims to solve the pessimistic value. Constraints (11) and (12) ensure that the conditions hold at confidence levels \( \beta_1 \) and \( \beta_2 \).

There are many uncertain variables in the above models. To solve the two models, the uncertainty inverse distribution technique in accordance with the uncertainty theory is to be introduced and discussed in the next section.

\section{Equivalence proof}

In many uncertain programming literatures pieces, various optimization methods are used to find an approximate optimal solution. The following will demonstrate that the two uncertain models can be converted into deterministic forms.

\textbf{Theorem 2.} The expected value model is equivalent to the following model:

\[
\min \left\{ m \sum_{i \in I} \sum_{l \in L} H_{jl} \sum_{i \in I} x_{ijl} + m \sum_{i \in I} \sum_{j \in J} \sum_{l \in L} C_{ijl} x_{ijl} + \sum_{j \in J} F_{j}(\sum_{i \in I} \sum_{l \in L} \sum_{j \in J} w_{ijl}) + m \sum_{i \in I} \sum_{j \in J} \sum_{l \in L} r_{i}^{j} S_{jl} \right\} \leq \bar{f} \tag{18}
\]

subject to:

\[
m \sum_{i \in I} x_{ijl} \leq \int_{0}^{1} \Phi_{A_{il}}^{-1}(\alpha) d\alpha, \quad l \in L, \tag{19}
\]

\[
m \sum_{i \in I} \sum_{j \in J} z_{jk} \int_{0}^{1} \Phi_{B_{jl}}^{-1}(\alpha) d\alpha, \quad j \in J, \tag{20}
\]

\[
\sum_{i \in I} \sum_{l \in L} w_{ijl} - M_{jl} y_{j} \leq 0, \quad j \in J \tag{21}
\]

\[
\sum_{j \in J} y_{j} - P \leq 0, \tag{22}
\]

\[
\sum_{j \in J} z_{jk} = 1, \quad k \in K, \tag{23}
\]

\[
\sum_{k} z_{jk} - r y_{j} \leq 0, \quad j \in J, \tag{24}
\]

\[
x_{ijl} \geq 0, \quad y_{j} = \{0, 1\}, \quad z_{jk} = \{0, 1\}, \tag{25}
\]

where \( \alpha, \beta_1, \beta_2 \) are the preset confidence levels.

\textbf{Proof:} According to the nature of expected value, the conclusion is easy to draw.

\[
E \left[ \sum_{j \in J} \sum_{k \in K} B_{jk} z_{jk} \left( \sum_{l \in L} q_{l} D_{kl} \right) \right] \]

\[
= \sum_{j \in J} \sum_{k \in K} \sum_{l \in L} z_{jk} q_{l} E[B_{jk} D_{kl}]. \]
Assume that the uncertain distributions of $B_{jk}$ and $D_{kl}$ are $\Phi_{B_{jk}}(x)$ and $\Phi_{D_{kl}}(x)$, respectively. According to Theorem 1, it yields:

$$E[B_{jk}D_{kl}] = \int_0^1 \Phi_{B_{jk}}^{-1}(\alpha)\Phi_{D_{kl}}^{-1}(\alpha)d\alpha,$$

because:

$$E[A_d] = \int_0^1 \Phi_{A_{d1}}^{-1}(\alpha)d\alpha,$$

then, Inequation (18) is equivalent to:

$$m \sum_{j \in J} x_{ijl} \leq \int_0^1 \Phi_{A_{d1}}^{-1}(\alpha)d\alpha.$$

Likewise, the equivalent forms of other constraints can be obtained.

The theorem is proved.

**Theorem 3.** The chance-constrained model is equivalent to the following model:

$$ \begin{align*}
\min & \quad m \sum_{i \in I} \sum_{l \in L} H_{ij} \sum_{j \in J} x_{ijl} + m \sum_{i \in I} \sum_{j \in J} \sum_{l \in L} C_{ijl}x_{ijl} \\
& \quad + \sum_{j \in J} F_j \left( \sum_{i \in I} \sum_{l \in L} w_{ijkl} x_{ijl} \right) m \sum_{i \in I} \sum_{j \in J} \sum_{l \in L} \sum_{\tau=1}^n \\
& \quad U_{ij} \frac{x_{ijl}}{n} + m \sum_{j \in J} \sum_{i \in I} \left( \sum_{l \in L} x_{ijl} \right) S_{ijl} \\
& \quad + \sum_{j \in J} \sum_{k \in K} \sum_{l \in L} z_{jk} q_{kl} \Phi_{B_{jk}}^{-1}(\alpha) \Phi_{D_{kl}}^{-1}(\alpha),
\end{align*}$$

subject to:

$$m \sum_{j \in J} x_{ijl} \leq \Phi_{A_{d1}}^{-1}(1 - \beta_1), \ i \in I, l \in L,$$

$$m \sum_{i \in I} x_{ijl} = \sum_{k \in K} z_{jk} \Phi_{D_{kl}}^{-1}(\beta_2), \ j \in J, l \in L,$$

$$\sum_{i \in I} \sum_{l \in L} w_{ijkl} - M_{ij} y_{ij} \leq 0, \ j \in J,$$

$$\sum_{j \in J} y_{ij} - P \leq 0,$$

$$\sum_{j \in J} z_{jk} = 1, \ k \in K,$$

$$\sum_{k} z_{jk} - r y_{ij} \leq 0, \ j \in J,$$

$$x_{ijl} \geq 0, \ y_{ij} = \{0, 1\}, \ z_{jk} = \{0, 1\}, \ i \in I,$$

$$j \in J, l \in L,$$

where $\Phi_{f}^{-1}$ denotes the inverse uncertainty distribution of $f$.

**Proof:** According to Definition 1, Inequations (9) and (10) are equivalent to Inequation (19).

According to the definition of uncertain distribution, it yields:

$$\beta_1 \leq \mathcal{M} \left\{ m \sum_{j \in J} x_{ijl} \leq A_{d} \right\}$$

$$= 1 - \mathcal{M} \left\{ m \sum_{j \in J} x_{ijl} > A_{d} \right\}$$

$$= 1 - \Phi \left( m \sum_{j \in J} x_{ijl} \right).$$

By taking inverse distribution on both sides, it yields:

$$m \sum_{j \in J} x_{ijl} \leq \Phi_{A_{d1}}^{-1}(1 - \beta_1).$$

Likewise, the equivalent forms of other constraints can be obtained.

The theorem is proved.

It has always been known that because of the multiple types of uncertainty, policymakers will face the problem of multi-dimensional decision variables. These variables lead to multiple integration problems in a random environment, thereby making the calculation more difficult to achieve. Fortunately, the problem of multiple integration is avoided by the operation law of inverse uncertainty distribution. Thus, the proposed uncertainty model outperforms the stochastic model in many types of uncertain facility location problems.

5. **Hybrid algorithm**

It is clear that the two deterministic models are nonlinear and NP-hard and they cannot be solved by exact methods [4]. Accordingly, it is necessary to find an effective algorithm to solve the deterministic forms of the models. Fortunately, meta-heuristic can effectively solve such complex problems, e.g., GA and PSO. Jiang et al. [40] proposed an effective method called GAP-SO-I to solve the distribution problem in B2C e-commerce. Inspired by the mentioned process, an improved GA was proposed according to the characteristics of the uncertain model. The Hybrid Algorithm (HA) looks for optimal costs among DCs, customers, and suppliers. The proposed algorithm HA is given in the following section.

- Solution representation and initialization: The distribution between the DC and the customer is
represented by a natural number. For instance, there are 4 potential DCs and 6 customers. A maximum of four DCs can be selected and the code can be written as [3, 1, 2, 2, 1, 4]. The column and element represent the customer and DC, respectively. This code dictates that DC 3 services customer 1, DC 1 services customer 2, DC 2 services customer 3, and the rest can be deduced by a similar analogy. To ensure that each DC is properly selected, four different codes should appear in the code. Four numbers 1, 2, 3, 4 are randomly arranged at six locations and two numbers are randomly generated at the other two locations. Once the initial population is determined, the values of \( y_j \) and \( z_{jk} \) will be determined.

- **Fitness function:** Use the objective function as a fitness function:

\[
f_E = \sum_{j \in J} \sum_{k \in K} \sum_{l \in L} z_{jkl} E[B_{jkl}],
\]

\[
f_P = \sum_{j \in J} \sum_{k \in K} \sum_{l \in L} z_{jkl} \Phi^{-1}_{B_{jkl}}(\alpha) \Phi^{-1}_{D_{jkl}}(\alpha).
\]

For the remaining sub-models:

\[
\min G(x_{ijl}),
\]

subject to:

\[
m \sum_{j \in J} x_{ijl} \leq \int_0^1 \Phi^{-1}_{A_{il}}(\alpha) d\alpha, i \in I, l \in L,
\]

\[
m \sum_{i \in I} x_{ijl} = \sum_{k \in K} z_{jkl} \int_0^1 \Phi^{-1}_{D_{jkl}}(\alpha) d\alpha, j \in J, l \in L,
\]

\[
\sum_{i \in I} \sum_{j \in J} w_{ij} x_{ijl} - M_j y_j \leq 0, j \in J,
\]

\[
x_{ijl} \geq 0, i \in I, j \in J, l \in L,
\]

\[
\min G(x_{ijl})
\]

subject to:

\[
m \sum_{j \in J} x_{ijl} \leq \Phi^{-1}_{A_{il}}(1 - \beta_1), i \in I, l \in L,
\]

\[
m \sum_{i \in I} x_{ijl} = \sum_{k \in K} z_{jkl} \Phi^{-1}_{D_{jkl}}(\beta_2), j \in J, l \in L,
\]

\[
\sum_{i \in I} \sum_{j \in J} w_{ij} x_{ijl} - M_j y_j \leq 0, j \in J,
\]

\[
x_{ijl} \geq 0, i \in I, j \in J, l \in L,
\]

\[
G(x_{ij}) = m \sum_{i \in I} \sum_{l \in L} H_{il} \sum_{j \in J} x_{ij} + m \sum_{i \in I} \sum_{j \in J} \sum_{l \in L} C_{ijl} x_{ij} \]

\[
+ \sum_{j \in J} F_j \left( \sum_{i \in I} \sum_{l \in L} w_{ij} x_{ij} \right) n
\]

\[
\sum_{i \in I} \sum_{j \in J} \sum_{l \in L} \left( \sum_{i \in I} x_{ijl} \right)^\theta S_j.
\]

Obviously, only decision variables \( x_{ijl} \) are covered in the sub-model. Since \( G(x_{ijl}) \) is nonlinear, conventional algorithms (e.g., branch and bound) cannot solve this problem well. PSO has aroused increasing attention for its easy implementation in recent years. A PSO is proposed to solve the sub-model as follows.

The position of the \( i \)th particle is denoted by \( X_i = (X_{i1}, X_{i2}, \ldots, X_{idm}) \), which is used to represent the three-dimensional subscript variable \( x_{ijl} \), and the velocity is denoted by \( V_i = (V_{i1}, V_{i2}, \ldots, V_{idm}) \). Let \( P^\text{best}_i \) and \( G^\text{best} \) be the local and global extrema, respectively. The inertia weight is \( w \). The cognition coefficient is \( c_1 \) and the social coefficient is \( c_2 \), and \( rand_1, rand_2 \sim U(0, 1) \).

The update formula is as follows:

\[
V_i(t+1) = w V_i(t) + c_1 \times rand_1(0, 1) \times (P^\text{best}_i - X_i(t)) + c_2 \times rand_2(0, 1) \times (G^\text{best} - X_i(t)),
\]

\[
X_i(t + 1) = X_i(t) + V_i(t + 1).
\]

If a criterion is met, stop; otherwise, perform another iteration.

- **Selection operator:** Selection process is based on the evaluation function of the population (the roulette wheel selection).

- **Crossover process:** Crossover is the process of producing offspring. To search solution space more fully, cross is used to produce better offspring. Two-point crossover method is used. The crossover probability is \( p_c \in (0, 1) \). Two cutting points are randomly assigned. Genes beyond the cutting points in parents 1 and 2 are directly duplicated.
to the offspring. Figure 1 shows the crossover operation.

However, infeasible solutions can arise from the crossover process. For instance, the number of DCs is 4. Figure 1 is an example of an infeasible solution. The solution does not satisfy Eq. (5). In this case, 6 DCs are selected. The repair method is as follows:

**Step 1**: The number of DCs is denoted by $Q$.

**Step 2**: If $Q \leq P$, the solution is feasible; otherwise, if $Q > P$, $Q = Q - P$.

**Step 3**: Select two different genes from the chromosomes randomly and then, let them take the same value. For instance, among the solutions $[3, 1, 2, 4, 3, 6, 1, 4, 2, 5]$, select 3 and 1 and then, turn them into $[3, 3, 2, 4, 3, 6, 3, 4, 2, 5]$ or $[1, 1, 2, 4, 1, 6, 1, 4, 2, 5]$. Repeat this iterative process $Q$ times until a feasible solution can be obtained.

**Mutation process**: The function of the mutation operator is to avoid the premature from falling into the local optimal. Inversion mutation is used to produce a feasible solution. The method is given in Figure 2. The reversal mutation is accomplished by randomly selecting two positions and the genes between the two locations are reversed.

**Termination**: If the maximum number of iterations is reached, stop; otherwise, circulate the selection process.

**6. Numerical experiment**

Some numerical examples are given in this section.

Assume that there are 4 suppliers, 6 potential DCs, 14 customers, and two types of commodities. Assume that there are 4 suppliers, 6 potential DCs, 14 customers, and two types of commodities.

$$w_1 = 0.7, \quad w_2 = 0.5, \quad q_1 = 100, \quad q_2 = 90,$$

$$m = 18, \quad n = 12, \quad P = 5, \quad \theta = 0.5,$$

$$E_{j0} = 12000, \quad \alpha = \beta_1 = \beta_2 = 0.9.$$

Other parameters are randomly generated (Tables 1 and 2). Assume that all uncertain variables follow the zigzag distribution.

The error of objective value often serves as a crucial tool to assess the merits of algorithms. Its expression:

$$\text{Error} = \frac{\text{objective value} - \text{the optimal value}}{\text{the optimal value}} \times 100\%.$$

The robustness of the HA is tested with GA and PSO under different parameters. First, the chance-constrained model is tested. The results are listed in Table 3.

**Table 1. Parameters from suppliers to Distribution Centers (DCs).**

<table>
<thead>
<tr>
<th>Supplier</th>
<th>DC</th>
<th>$C_{ijt}$</th>
<th>$U_{ij}$</th>
<th>$S_{ij}$</th>
<th>$M_{ij}$</th>
<th>$N_{ij}$</th>
<th>$F_{j0}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>10</td>
<td>16</td>
<td>15</td>
<td>13</td>
<td>12</td>
<td>40</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>8</td>
<td>13</td>
<td>14</td>
<td>16</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>11</td>
<td>14</td>
<td>9</td>
<td>15</td>
<td>10</td>
<td>105</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>9</td>
<td>17</td>
<td>14</td>
<td>9</td>
<td>12</td>
<td>110</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>16</td>
<td>8</td>
<td>10</td>
<td>11</td>
<td>9</td>
<td>13</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>15</td>
<td>10</td>
<td>13</td>
<td>18</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13</td>
<td>12</td>
<td>9</td>
<td>18</td>
<td>15</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12</td>
<td>8</td>
<td>10</td>
<td>13</td>
<td>9</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3000</td>
<td>2800</td>
<td>3400</td>
<td>3200</td>
<td>3500</td>
<td>3600</td>
</tr>
</tbody>
</table>
Table 2. Unit delivery cost from Distribution Centers (DCs) to customers.

<table>
<thead>
<tr>
<th>$B_{jk}$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(4.5,6)</td>
<td>(3.4,5)</td>
<td>(7.8,9)</td>
<td>(6.7,8)</td>
<td>(8.9,10)</td>
<td>(11.12,13)</td>
<td>(8.9,10)</td>
</tr>
<tr>
<td>2</td>
<td>(3.4,5)</td>
<td>(7.8,9)</td>
<td>(8.9,10)</td>
<td>(3.4,5)</td>
<td>(11.12,13)</td>
<td>(7.8,9)</td>
<td>(4.5,6)</td>
</tr>
<tr>
<td>3</td>
<td>(7.8,9)</td>
<td>(3.4,5)</td>
<td>(4.5,6)</td>
<td>(2.3,4)</td>
<td>(8.9,10)</td>
<td>(11.12,13)</td>
<td>(6.7,8)</td>
</tr>
<tr>
<td>4</td>
<td>(3.4,5)</td>
<td>(8.9,10)</td>
<td>(6.7,8)</td>
<td>(11.12,13)</td>
<td>(7.8,9)</td>
<td>(2.3,4)</td>
<td>(7.8,9)</td>
</tr>
<tr>
<td>5</td>
<td>(2.3,4)</td>
<td>(4.5,6)</td>
<td>(4.5,6)</td>
<td>(3.4,5)</td>
<td>(8.9,10)</td>
<td>(9.10,11)</td>
<td>(3.4,5)</td>
</tr>
<tr>
<td>6</td>
<td>(7.8,9)</td>
<td>(7.8,9)</td>
<td>(2.3,4)</td>
<td>(4.5,6)</td>
<td>(6.7,8)</td>
<td>(6.7,8)</td>
<td>(9.10,11)</td>
</tr>
</tbody>
</table>

Table 3. Solutions under different parameters (chance-constrained model).

<table>
<thead>
<tr>
<th>GA</th>
<th>No.</th>
<th>Size</th>
<th>max_ite</th>
<th>$p_c$</th>
<th>$p_m$</th>
<th>PSO</th>
<th>Optimal solution</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>20</td>
<td>1000</td>
<td>0.8</td>
<td>0.15</td>
<td>20</td>
<td>30</td>
<td>16941223.79</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>20</td>
<td>1000</td>
<td>0.75</td>
<td>0.2</td>
<td>15</td>
<td>40</td>
<td>16950746.51</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>20</td>
<td>1000</td>
<td>0.7</td>
<td>0.25</td>
<td>10</td>
<td>50</td>
<td>16988996.17</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>30</td>
<td>800</td>
<td>0.8</td>
<td>0.15</td>
<td>20</td>
<td>30</td>
<td>16939121.63</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>30</td>
<td>800</td>
<td>0.75</td>
<td>0.2</td>
<td>15</td>
<td>40</td>
<td>16955874.32</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>30</td>
<td>800</td>
<td>0.7</td>
<td>0.25</td>
<td>10</td>
<td>50</td>
<td>16972242.36</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>40</td>
<td>500</td>
<td>0.8</td>
<td>0.15</td>
<td>20</td>
<td>30</td>
<td>16945951.05</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>40</td>
<td>500</td>
<td>0.75</td>
<td>0.2</td>
<td>15</td>
<td>40</td>
<td>16962013.18</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>40</td>
<td>500</td>
<td>0.7</td>
<td>0.25</td>
<td>10</td>
<td>50</td>
<td>16989477.34</td>
</tr>
</tbody>
</table>

The errors are not larger than 0.08795 under different parameters. Besides, the mean value is 16961290.7 and the average error is 0.037204. This suggests that the changes in the parameters slightly affect the optimal value and it is therefore indicated that the proposed algorithm exhibits good robustness. In contrast, the optimal value is in the fourth order, as shown in Table 3.

GAPSO [40] is also an effective algorithm for this kind of problem. Two measures (objective value and CPU time) are used to assess the effectiveness of the HA and GAPSO. The comparative results are listed in Tables 4 and 5. According to the results in two tables, HA generally outperforms GAPSO. The CPU time of HA is slightly less than that of the GAPSO. The CPU times of the two algorithms are perfectly acceptable in practice. Moreover, the objective values vary slightly with the modeling angle.

To study the sensitivity of $\alpha$, $\beta_i$, $i = 1, 2$ in the chance-constrained model, another supplementary test is performed and the results are listed in Figure 3. The step size of the confidence level is taken as 0.2. Figure 3 implies the objective function is nondecreasing with $\alpha$, $\beta_i$, $i = 1, 2$. The result of the sensitivity analysis allows decision-makers to make most reasonable judgment based on the degree of understanding of actual problems in an uncertain environment.
Table 4. Results of the expected value model

<table>
<thead>
<tr>
<th>No.</th>
<th>GAPSO Objective value</th>
<th>GAPSO CPU time (s)</th>
<th>HA Objective value</th>
<th>HA CPU time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>16805171.59</td>
<td>232.55</td>
<td>16290511.19</td>
<td>232.89</td>
</tr>
<tr>
<td>2</td>
<td>16882147.82</td>
<td>233.69</td>
<td>16205177.07</td>
<td>234.06</td>
</tr>
<tr>
<td>3</td>
<td>16902643.66</td>
<td>231.74</td>
<td>16250213.49</td>
<td>232.22</td>
</tr>
<tr>
<td>4</td>
<td>16930144.03</td>
<td>233.79</td>
<td>16159885.09</td>
<td>231.14</td>
</tr>
<tr>
<td>5</td>
<td>16920465.18</td>
<td>231.35</td>
<td>16240957.26</td>
<td>230.01</td>
</tr>
<tr>
<td>6</td>
<td>16901315.23</td>
<td>232.91</td>
<td>16300258.66</td>
<td>231.45</td>
</tr>
<tr>
<td>7</td>
<td>16963331.57</td>
<td>231.18</td>
<td>16225510.15</td>
<td>230.97</td>
</tr>
<tr>
<td>8</td>
<td>16885559.01</td>
<td>232.08</td>
<td>16313300.53</td>
<td>231.63</td>
</tr>
<tr>
<td>9</td>
<td>16957885.38</td>
<td>231.89</td>
<td>16302971.88</td>
<td>230.04</td>
</tr>
<tr>
<td>10</td>
<td>16821659.97</td>
<td>232.86</td>
<td>16298821.51</td>
<td>231.43</td>
</tr>
<tr>
<td>Average</td>
<td>16906632.4</td>
<td>232.4</td>
<td>16258703.7</td>
<td>231.58</td>
</tr>
</tbody>
</table>

Table 5. Comparative results of Hybrid Algorithm (HA) and Genetic Algorithm and Particle Swarm Optimization (GAPSO) (chance-constrained model).

<table>
<thead>
<tr>
<th>No.</th>
<th>GAPSO Objective value</th>
<th>GAPSO CPU time (s)</th>
<th>HA Objective value</th>
<th>HA CPU time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>16473861.13</td>
<td>235.06</td>
<td>16098545.97</td>
<td>231.33</td>
</tr>
<tr>
<td>2</td>
<td>16885573.54</td>
<td>231.14</td>
<td>15962665.22</td>
<td>232.94</td>
</tr>
<tr>
<td>3</td>
<td>16634665.71</td>
<td>229.97</td>
<td>16055433.27</td>
<td>230.87</td>
</tr>
<tr>
<td>4</td>
<td>16851349.55</td>
<td>232.05</td>
<td>16145333.63</td>
<td>230.55</td>
</tr>
<tr>
<td>5</td>
<td>16512146.03</td>
<td>234.93</td>
<td>16090876.34</td>
<td>231.21</td>
</tr>
<tr>
<td>6</td>
<td>16659617.11</td>
<td>234.56</td>
<td>15923112.46</td>
<td>232.96</td>
</tr>
<tr>
<td>7</td>
<td>16767988.05</td>
<td>233.28</td>
<td>16025510.76</td>
<td>233.04</td>
</tr>
<tr>
<td>8</td>
<td>16821138.21</td>
<td>235.77</td>
<td>16122908.12</td>
<td>231.53</td>
</tr>
<tr>
<td>9</td>
<td>16457855.64</td>
<td>232.09</td>
<td>16021676.96</td>
<td>231.14</td>
</tr>
<tr>
<td>10</td>
<td>16524975.34</td>
<td>234.63</td>
<td>15898821.03</td>
<td>230.66</td>
</tr>
<tr>
<td>Average</td>
<td>16641920</td>
<td>233.35</td>
<td>16034797.4</td>
<td>231.62</td>
</tr>
</tbody>
</table>

Table 6. Large-scale problem (chance-constrained model).

<table>
<thead>
<tr>
<th>No.</th>
<th>r</th>
<th>GAPSO Best</th>
<th>GAPSO Ave</th>
<th>GAPSO Worst</th>
<th>GAPSO Time (s)</th>
<th>HA Best</th>
<th>HA Ave</th>
<th>HA Worst</th>
<th>HA Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50</td>
<td>98.99</td>
<td>94.57</td>
<td>91.96</td>
<td>2159.02</td>
<td>100</td>
<td>99.08</td>
<td>97.68</td>
<td>2133.41</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>97.95</td>
<td>95.03</td>
<td>92.15</td>
<td>2330.54</td>
<td>100</td>
<td>98.92</td>
<td>97.55</td>
<td>2219.12</td>
</tr>
<tr>
<td>3</td>
<td>200</td>
<td>99.04</td>
<td>95.37</td>
<td>92.19</td>
<td>2192.22</td>
<td>100</td>
<td>98.87</td>
<td>97.66</td>
<td>2178.93</td>
</tr>
<tr>
<td>4</td>
<td>300</td>
<td>99.36</td>
<td>96.92</td>
<td>92.30</td>
<td>2260.17</td>
<td>100</td>
<td>98.29</td>
<td>97.54</td>
<td>2239.11</td>
</tr>
<tr>
<td>5</td>
<td>400</td>
<td>99.21</td>
<td>96.82</td>
<td>92.63</td>
<td>2204.56</td>
<td>100</td>
<td>99.27</td>
<td>97.09</td>
<td>2186.36</td>
</tr>
<tr>
<td>6</td>
<td>500</td>
<td>99.52</td>
<td>96.43</td>
<td>91.98</td>
<td>2229.31</td>
<td>100</td>
<td>99.13</td>
<td>98.95</td>
<td>2208.94</td>
</tr>
</tbody>
</table>

To test the proposed algorithm further, the problem is investigated on a large scale. The uncertain demands are generated randomly. Since the problem is more complex on a large scale, the parameters are set as follows:

c = 250, max_ite = 500.

Other parameters do not adjust. To avoid random large errors random, two algorithms are tested 10 times, respectively. For the sake of comparison, we have:

quality of objective value = \frac{the~optimal~value}{the~current~value} \times 100\%.

The results of the two algorithms are listed in Table 6. The result of the expected value model is similar to that of the chance-constrained model, and it is omitted. According to the results, these two algorithms can still solve such problems effectively on a large scale.
In the above examples, the expected value model and chance-constrained model are used to solve the supply chain problem in an uncertain environment. It is noteworthy that two different styles of models under different guidelines are proposed in this study. According to the results of the two examples, there is a relative difference between the two solutions of the two uncertain models, primarily because the two models are built from different perspectives, thereby resulting in different optimal solutions. It is difficult to generalize which model is better. In fact, a more suitable model is determined by the decision-maker and the mastery of the actual situation.

7. Conclusions

The problem of an uncertain facility location in Business-to-Consumer (B2C) e-commerce was investigated in this study. Unlike the past, delivery cost as well as supply and demand were assumed to be uncertain variables due to the lack of observed data. To deal with these empirical data, the expected value model and chance-constrained model were developed. To overcome the limitation of the capacity of the Distribution Center (DC), a more reasonable cost function of DCs was established. The equivalent forms of these models were obtained in accordance with the uncertainty theory. An improved Genetic Algorithm (GA) with Particle Swarm Optimization (PSO) was proposed to find an optimal approximate solution. The effectiveness and efficiency of the proposed models were verified by several numerical experiments. Besides, according to the results of the extensive computational experiments, the proposed hybrid algorithm is more competitive and efficient than GAPSO. Furthermore, this modeling idea and solution method may also be suitable for solving other facility location problems.

Acknowledgments

I am grateful to the editor and the anonymous reviewers for their helpful suggestions on an earlier version of this paper. This work is supported by the National Natural Science Foundation of China (No. 61673011) and Research Foundation of NIT (YK18-10-02, YK18-10-03).

References


Biography

Jiaiyu Shen received the BS degree in information and computing science from Hangzhou Dianzi University, Hangzhou, China in 2007 and the PhD degree in mathematics from Nanjing University of Science and Technology, Nanjing, China in 2016. He is now a Lecturer of the Department of Public Basic Courses at Nanjing Institute of Industry Technology. His research interests include uncertain scheduling theory, supply chain management and optimal control.