Research Article

Integrated Data Mining and TOPSIS Entropy Weight Method to Evaluate Logistics Supply and Demand Efficiency of a 3PL Company

Xintong Yan, Jian Gong, Jie He, Hao Zhang, Changjian Zhang, and Ziyang Liu

School of Transportation, Southeast University, 2 Si Pai Lou, Nanjing 210096, China

Correspondence should be addressed to Jie He; hejie@seu.edu.cn

Received 13 December 2019; Revised 26 February 2020; Accepted 10 March 2020; Published 2 April 2020

Copyright © 2020 Xintong Yan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the past decades, despite considerable attention having been paid to third-party logistics (3PL) owing to its specialized service, sophisticated operation, and reduced cost, research on quantitative methods for estimating the efficiency of 3PL companies is still lacking, especially for those with small or medium scale. Therefore, the purpose of this study was to establish a quantitative evaluation method to measure the efficiency of the individual nongovernmental 3PL firms and explore the valuable information for the management of 3PL business with Apriori and K-means. Taking TopChains (an emerging nongovernmental 3PL company) as an example, the monthly supply and demand (S&D) level and matching degree were evaluated via the integrating data mining algorithm (i.e., Apriori and K-means) and TOPSIS entropy weight method based on historical data. The findings demonstrate that the S&D level varied with time and space, and the customer demand in February tended to reduce substantially. Besides, the outcome of S&D matching degree in June is undesirable, indicating the unsatisfactory efficiency in resource management. The evaluation maneuver stated in this study can serve as a valuable tool to measure individual nongovernmental 3PL enterprises’ efficiency in terms of S&D, and for reference, the results can aid in rational enterprise investment plan. Besides, this attempt broadened the direction of ARM and K-means being applied in the logistics field.

1. Introduction

Both the trend of economic globalization and the evolution of information technology have expanded the scope of production, sales, and circulation of commodities, which generates a larger logistics market, requiring more specialized and informative logistics service. The third-party logistics (3PL), launched in the U.S. and Europe in the 1980s and focusing on the sophisticated logistics operation mode, can reduce the operating costs and enhance the service level, thus rapidly occupying the logistic market [1–5].

1.1. 3PL Capacity and Efficiency. Despite the growing trend of 3PL with the expectation to reduce the total costs through efficient utilization, the service is not completely efficient [5, 6]. Scholars have grasped the phenomenon and focused on the subjects about 3PL capacity and efficiency [7–9].

The capacity of logistics is generally summarized as the capacity to satisfy the needs of circulation of materials, associated with various processes, e.g., transporting, storage, and delivery, which can be measured by the relationship of the supply and demand. When supply exceeds demand, it will cause a waste of resources, which is inefficient. On the contrary, if supply is overwhelmed by demand, there will be an undesirable declination of the quality of logistics services, also inducing the lack of efficiency in customer service. Marchet et al. [5] have proposed that the current capability of 3PL is not sufficient enough to support existing even future needs in some special industries. Besides, due to the low industry threshold, many companies have entered into the 3PL market, leading to a fiercely competitive surrounding, while the overall capacity...
does not obtain a prominent amelioration for the lack of the efficiency.

Obviously, it is important for the 3PL industry to enhance its overall capacity, which can be realized through the improvement of the efficiency of the individual 3PL companies, among which the management of the relationship of supply and demand should be highlighted. Notably, almost each modern 3PL supplier is confronted with the pressure brought by both customers and their competitors [7, 10, 11] since the former requires the sophisticated services while the latter can be regarded as one of the biggest external threats, especially for some emerging nongovernmental 3PL companies with small or medium size.

Given the significance of the enhancement of the efficiency of individual 3PL companies to acquire higher profits with lower cost, an exploration of evaluating the efficiency of individual 3PL companies is warranted to help individual 3PL enterprises to better assess their performance and thus increase the efficiency.

1.2. Relationship of Supply and Demand. The empirical evidence gathered from logistics practice demonstrates that the logistics efficiency can be influenced by the relationship of the logistics supply and demand, and the balance of the logistics S&D has a positive impact on the efficiency of 3PL enterprises. However, the management of the relationship between the supply and demand is still a challenge due to the complexity caused by the many unpredictable factors [12, 13]. Logistics supply is of uncertainty due to the changes and the unpredictability of logistics providers’ prices, capabilities, quality, and so on [2]. Similarly, logistics demand tends to be uncontrollable and fluctuant, influenced by not only prices but also season, distribution, and service. Both supply and demand uncertainty could trigger risks for 3PL enterprises, especially for those budget-constrained ones [14].

The quantitative evaluation of the logistics supply and demand can provide a powerful instrument for the 3PL organizations to measure their efficiency. The existing research generally analyzes supply and demand (S&D) separately; it is worth noting that the connection between the supply and demand should be attached importance too. Some scholars did explore the interaction of logistics S&D, for example, Ottemöller and Friedrich [15] have presumed the conversion of freight demand might be consequent upon the transformation of the supply chain structure and thus constructed a potential factor model to demonstrate their assumption from the perspective of the reasonable allocation of the commodity flow and equipment flow. But, for some constraints such as low data availability, uncertainty issues, and heavy calculation burden, the relevant quantification systems of evaluating the relationship of S&D are still immature.

1.3. Traditional 3PL Quantitative Evaluation Models. In the field of logistics, many quantitative models have been established to explore the complexity of the supply chain [16–21]. Some quantitative methods have been employed for the assessment of the operational capacity and efficiency of 3PL with the purpose of selecting the reliable suppliers as cooperative partners and providing strategic basis for defining the policy [1, 5, 13, 22, 23]. Owing to the considerable significance of assessing and thus improving the performance of the 3PL, the application of quantitative models in dealing with the 3PL issues tends to be a promising topic in the logistics subject. Besides, various factors can be involved when constructing the 3PL evaluation model such as management success (e.g., organizing ability, input-output efficiency, and equipment utilization), business strength (e.g., financial situation and representative performance), service quality (e.g., information construction situation, information receiving and processing rate, and customer satisfaction), and business growth (e.g., enterprise scale and management concept), and they can be adjusted combined with the practical selection issues [24].

The most prevailing method for 3PL supplier assessment is AHP, for example, to assess the service capacity of 3PL enterprises, Ecer [1] used AHP to establish an evaluation model based on Distance from Average Solution pertaining to three features of the 3PL suppliers: specialty, cost, and quality; Shakourloo et al. [22] proposed a more effective supplier selection framework based on AHP factors to specify the priority of products for the allocation to the supplier considering the product profit, and a method for assessing the efficiency of shipping has also been set up by Chen [25] based on the AHP. Besides these, AHP can be combined with other methods to realize the optimal ranking and select the suppliers [26], e.g., Jovčić et al. [27] conducted an integration of AHP and TOPSIS for a 3PL supplier evaluation model, so did Akman and Baynal [28]; Jayant et al. [29]; and Sun [30].

Despite the fact that AHP has been illustrated as a useful instrument to evaluate the performance of the 3PL suppliers, there are still some limitations in practical dimension, e.g., the decision-making process is sometimes subjective, leading to the unreliability and inauthenticity of the evaluation results, and the requirement for the collection of different types of data is unrealistic, expensive, and time consuming for some small-sized and medium-sized 3PL supplier, making it less successful to establish an efficient evaluation method. For some emerging 3PL companies, the ability to process data concisely and quickly is of great necessity, which implies that further research is still required to find new quantitative evaluation methods of robust and powerful data processing capacity.

1.4. Data Mining Algorithm. Solving the aforementioned issues requires advanced tools and techniques, and data mining (DM) could be regarded as an optimal algorithm to find potential information and extract 3PL S&D data in batchwise, real time, and near-time owing to its advanced analytic capability to make efficient, intelligent, and timely decisions [31, 32]. DM is roughly classified into six categories: classification, regression, clustering, prediction, association, and diagnosis, among which the most
representative is C4.5, K-means, SVM, Apriori, EM, PageRank, AdaBoost, KNN, Naive Bayes, and CART [33].

Some literature is conducted to reveal the priority of DM on the condition that proper algorithms are used to deal with the logistics problems [34, 35], while other scholars and institutions have applied the relevant data mining methods to the field of logistics, [36–39]. Most of the studies, however, concentrate on the area of monopolistic supply chain management, not specific enough, and the analyses of the S&D issue are always fragmented since the demand can be commonly postulated through the regression and prediction DM algorithm while the supply is hard to define and measure.

It is possible that other DM algorithms have an impressive capability to explore the significant information between logistics S&D. Li [40] has argued that the association rule mining (ARM) is technically feasible to deal with the business data and find new insights, while the relevant research still lacks. Besides, it seems that fuzzy clustering means (FCM) can serve as a valuable and fast tool to find the laws whilst few scholars have employed it to investigate the supply and demand efficiency of 3PL.

Therefore, the objectives of the current study are as follows: (1) to establish a feasible quantitative evaluation method to assess the efficiency of emerging nongovernmental 3PL companies measured by the relationship of S&D; (2) to explore the valuable information for the management of 3PL business; and (3) to justify that K-means and ARM could also serve as effective algorithms for the exploration of the efficiency of 3PL in terms of the relationship of the S&D.

2. Method

2.1. Data Acquisition. TopChains, an emerging nongovernmental 3PL enterprise founded in 2007 with 10 million yuan registered funds, has developed into a modern international logistics enterprise in possession of two branches in Shanghai and Shenzhen. Its business scale reaches 450 million yuan in 2018. At present, there are more than 450 employees and 120 transport vehicles, with a total warehouse area of more than 30,000 square meters. Compared with some large international 3PL enterprises, TopChains is a medium-sized enterprise, with its scale in the developing stage.

The supply of 3PL companies involves various factors (i.e., manpower, vehicles, warehouses, information systems, and other resources), increasing the difficulty in collecting the comprehensive data, especially for those who are small sized and medium sized, lacking advanced technique for the effective collection of various types of data. Xu and Zhang [41] have argued that the cost of 3PL companies can reflect the quality of customer service, which is generally associated with the supply of 3PL enterprises. Therefore, in order to reflect the level of supply more intuitively and concisely, the paper used the monthly cost of TopChains as the indicator to represent the level of supply. With regard to the level of demand, researchers have utilized the historical freight data to measure the demand of the logistics market, while hardly any research has evaluated the demand using the actual turnover data, which are more authentic since publicly available freight data are not reliable in many countries, including China. Besides, it is difficult for an emerging 3PL company to record an adequate number of freight volume over a short period of time to conduct an effective evaluation. Thus, the turnover data were utilized in this research to reflect the customer demand with the following merits: (1) the turnover data are easier to be collected compared to freight data and more realistic for an emerging 3PL company; (2) the turnover data provide more comprehensive understanding of the customer demand; and (3) the collection of turnover data is time saving and improves the efficiency.

The turnover and cost data of the branch companies in Shanghai and Shenzhen (from 2013 to 2018) are selected to measure the level of S&D of TopChains. To evaluate whether the operation of TopChains is efficient, the matching degree is defined to measure whether it is balanced between S&D. There are two phases for the evaluation approach: (1) the evaluation of S&D was conducted by K-means and Apriori, with K-means performing a preliminary judgement of monthly S&D to reflect the variances in time and space and then a detailed evaluation using Apriori undertaken to better manifest the S&D level and provide data for the next stage. (2) Based on the results of phase (1), the matching degree was calculated via the TOPSIS-entropy weight method.

2.2. Algorithm

2.2.1. K-Means Algorithm. K-means, as one of the common methods in the clustering algorithm, first proposed by MacQueen [42], has been widely used in the clustering algorithm for its effective and simple characteristics.

The K-means algorithm takes the distance between data as the standard of similarity measurement of data objects, so choosing the calculation method of distance between data has a significant impact on the final clustering effect. The commonly used methods of distance calculation are cosine distance, European distance, Manhattan distance, and so on. This paper takes Euclidean distance as a calculating standard because of its reliability and generality.

The European distance formula is as follows:

\[
\text{dist}(X_i, X_j) = \sqrt{\sum_{d=1}^{D} (x_{i,d} - x_{j,d})^2}. \tag{1}
\]

The distance between each pair of data objects can be calculated by formula (1).

According to the distance, the data objects can be clustered into a specified number of categories K. For each kind of data, the mean value of all data in the relative class is initially selected as the class center which needs iterating until the class center changes slowly or stops changing. Define the class center as

\[
\text{center}_k = \frac{1}{|C_k|} \sum_{X_i \in C_k} X_i, \tag{2}
\]

where \(C_k\) represents class \(k\) and \(|C_k|\) represents the number of data objects in class \(k\).
The iterating process can be calculated as
\[ J = \sum_{k=1}^{K} \sum_{X_c \in C_k} \text{dist}(X_c, \text{center}_k). \] (3)

2.2.2. Apriori Algorithm. Apriori algorithm is a representative algorithm for association rule mining (ARM), which is also one of the top ten data mining algorithms, originally proposed by Agrawal et al. [43], developing into an important research direction in data mining. Apriori has been widely used in the financial industry in many Western countries, with the main task to find out the internal relationship between things. However, in China, the research and application of Apriori is still not extensive enough, lacking of deep and valuable research. In this research, Apriori was utilized as a desirable instrument owing to its advantages of satisfactory data-processing capability and low requirements for the data types. Besides, Apriori is easier to understand and thus save the time and cost to process data, especially suitable for those budget-constrained enterprises.

ARM can be presented as the form of \( X \rightarrow Y \), where \( X \) and \( Y \) are disjoint item sets, i.e., \( X \cap Y = \emptyset \). The strength of ARM can be measured by its support and confidence. Support (s) can be used to determine the frequency of a given dataset, while confidence (c) determines the frequency of \( Y \) in transactions containing \( X \). The forms of support and confidence are defined as follows:
\[ s(X \rightarrow Y) = \frac{\sigma(X \cup Y)}{N}, \] (4)
\[ c(X \rightarrow Y) = \frac{\sigma(X \cup Y)}{\sigma(X)} \] (5)

The steps of Apriori are as follows:
(1) Let \( k = 1 \)
(2) Generate frequent item sets of length \( k \)
(3) Repeat until no new frequent itemsets are identified
(i) Generate length \((k + 1)\) candidate item sets from length \( k \) frequent item sets
(ii) Prune candidate item sets containing subsets of length \((k + 1)\) that are infrequent
(iii) Count the support of each candidate by scanning the database
(iv) Eliminate candidates that are infrequent, leaving only those that are frequent

2.2.3. TOPSIS Entropy Weight Method. At present, the commonly used evaluation methods include AHP and the fuzzy comprehensive evaluation method, but the subjective factors to be considered are various and complicated, which are not suitable for the analysis of this case. The combination of TOPSIS and the entropy weight method does not need subjective factors and variables, and the evaluation results are more objective and direct. Therefore, the TOPSIS entropy weight method is selected to evaluate the matching degree of logistics supply and demand.

TOPSIS (Technique for Order Preference by Similarity to an Ideal Solution) is a multiobjective decision method to determine the advantages and disadvantages of a scheme according to the weighted distance between the scheme and the positive and negative ideal values of indicators. Entropy weight method, as an evaluation method to determine the objective weight according to the variation of indicators, has been applied by some researchers on logistics management, for example, Li et al. [19] used the theory of entropy to analyze the channel operation and service of the supply chain. Based on the recovery rate, Li et al. [44] established another supply chain system with complexity using the theory of entropy.

The steps of the method can be demonstrated as follows:
(1) Use the entropy method to determine index weight:
Standardize the data according to
\[ Y_{ij} = \frac{x_{ij} - \min(x_i)}{\max(x_i) - \min(x_i)}. \] (6)
Calculate the information entropy
\[ E_j = -\frac{1}{\ln n} \sum_{i=1}^{n} p_{ij} \ln p_{ij}, \] (7)
where \( p_{ij} = Y_{ij}/\sum_{i=1}^{n} Y_{ij} \).
The weight of the matching index is calculated based on the formula
\[ W_i = \frac{1 - E_j}{k - \sum E_j}. \] (8)
(2) Use the TOPSIS algorithm to evaluate the target:
Establish a normalized weighted matrix
\[ Z_{ij} = W_i Y_{ij}. \] (9)
Determine the positive and negative ideal solution, and the positive ideal solution is
\[ L^+ = \begin{cases} \max Z_{ij} \ (J \text{ is a valid attribute}), \\
\min Z_{ij} \ (J \text{ is a cost attribute}), \end{cases} \] (10)
and the negative ideal solution is
\[ L^- = \begin{cases} \max Z_{ij} \ (J \text{ is a cost attribute}), \\
\min Z_{ij} \ (J \text{ is a valid attribute}). \end{cases} \] (11)
The distance from positive and negative ideal solutions to positive ideal solutions is
\[ S^+ = \sqrt{\sum_{j=1}^{m} (Z_{ij} - L^+_{j})^2}, \] (12)
and from negative ideal solutions is

\[ S' = \sqrt{\sum_{j=1}^{m} (Z_{ij} - L_j)^2}. \]  

(13)

The queuing index value can be calculated as

\[ Q_i = \frac{S_i}{S_i + S'_i}. \]  

(14)

2.3. Model

2.3.1. Monthly S&D Characteristic Analysis Based on K-Means. Based on the K-means algorithm, the monthly turnover and cost data from 2013 to 2018 of TopChains in Shanghai (TSH, CSH) and Shenzhen (TSZ, CSZ) were analyzed. The results were divided into 5 groups by a 5-level assessing scale ranging from 1 (the lowest) to 5 (the highest) according to the value of turnover and cost (shown in Table 1).

2.3.2. Establishment of Monthly S&D Evaluation Model Based on Apriori. The data in Section 2.3.1 were also used in this section. Besides, the total turnover and cost of two subsidiaries (TT and TC) were calculated to establish the initial data item set of ARM (as shown in Table 2).

For ARM, the required variables should be discrete, and it is necessary to transform the data into the feasible form. The data in Table 2 were discretized through ranking the discrete IDs ranges from 1 (with the highest value in a year) to 12 (the lowest) of each indicator, to estimate the monthly ranking ranging from 1 (the highest) to 12 (the value is the lowest), e.g., TSH in February 2013 is TSH12.

Apriori algorithm was carried out by SPSS Clementine. Month was defined as the preceding item of the ARM item set. The latter items were determined as TSH, TSZ, TT, CSH, CSZ, and TC relatively to obtain the result of ARM analysis (shown in Table 4).

This paper extracted the values of confidence of the six discrete variables calculated through ARM as indicators, and since the initial result is indirect to reflect the level of S&D, a linear model was designed, integrating the values of confidence and monthly ranking ranging from 1 (the highest) to 12 (the lowest) of each indicator, to estimate the monthly level of S&D.

The formula proposed in this paper is as follows:

\[ X_{pi} = \sum_{i=1}^{12} L_j \cdot a_p (i = 1, 2, \ldots, 6, j = 1, 2, \ldots, 12, p = \text{Jan, Feb, \ldots, Dec}). \]  

(15)

Among them, variables \( X_1, X_2, X_3, X_4, X_5, \) and \( X_6 \) shown in Table 4 were set up and scored. \( L_j \) is the score pertaining to the ranking, and from 1 to 6, the values of \( L_j \) are 6, 5, 4, 3, 2, and 1, while from 7 to 12, the values of \( L_j \) are \(-1, -2, -3, -4, -5, \) and \(-6; \) \( a_p \) is the value of confidence of each month (100%, 83.6%, 66.7%, 50%, 33.3%, 16.7%, and 0), which represents the frequency of the month appearing in the relative ranking. Both values of ranking and confidence are presented in Table 5.

Based on the historical data (2013–2018), the comprehensive scores of the level of S&D are shown in Table 6. It is worth noting that the higher score means the better value of the level of S&D.

2.3.3. Evaluation of S&D Matching Degree Based on the TOPSIS Entropy Weight Method

(1) Data Selection. The results of the monthly S&D level scoring in Table 6 of the association analysis are selected as the source data: \( X_1, X_2, X_3, X_4, X_5, \) and \( X_6, \) and the scoring items are \( Y_{p1}, Y_{p2}, \) and \( Y_{p3}. \)

The calculating formula is

\[ Y_{p1} = \begin{cases} \frac{X}{X_{p1}}, & X_{p1} < X_{p2}, \\ \frac{X}{X_{p2}}, & X_{p1} > X_{p2}, \end{cases} \]

\[ Y_{p2} = \begin{cases} \frac{X}{X_{p3}}, & X_{p3} < X_{p4}, \\ \frac{X}{X_{p4}}, & X_{p3} > X_{p4}, \end{cases} \]

\[ Y_{p3} = \begin{cases} \frac{X}{X_{p5}}, & X_{p5} < X_{p6}, \\ \frac{X}{X_{p6}}, & X_{p5} > X_{p6}, \end{cases} \]  

(16)

where \( p = \text{Jan, Feb, \ldots, Dec.} \)

The original matching degree score matrix is shown in Table 7.

(2) Calculation of Entropy Weight. Data are standardized according to formula (6). The results are shown in Table 8. According to formula (7), the information entropy

\[ E = (0.955, 0.948, 0.929). \]

The weight of the matching index is calculated based on formula (8), and the result is \( W = (0.266, 0.311, 0.422). \)

(3) TOPSIS Algorithm. According to formulas (8)–(14), programming with MATLAB, the final results of 12 months are shown in Table 9.

3. Results

3.1. Monthly S&D Level Analysis Based on K-Means. The result obtained in the first phase through the K-means algorithm illustrated that the level of demand and supply
showed no difference, and the overall S&D level (mean level = 2.25) in the first six months was lower than that (mean level = 4.15) in the second six of a year. It was demonstrated that the location of the company might have an impact on the time distribution of the S&D level (shown in Figures 1–4), as Zhang and Wang [45] have argued that time and location can be regarded as critical elements for the S&D network, and the S&D level of branches in Shanghai and Shenzhen proves divergent.

The main findings can be concluded as follows:

1. February is prone to be the lowest period of S&D level (mean = 1), which might be influenced by the Spring festival and extremely cold weather. Cold weather may have a negative impact on the logistics market; as China Logistics Times [46] has reported, the extremely cold weather in 2009 interfered with Haier’s supply chain and led to a reduction by 34.43% in its business.

2. The peak period for S&D level of Shanghai company is November (mean = 5), while that of Shenzhen Company is October to December (mean = 5). Besides, it is obvious that the S&D level in December of the branch in Shanghai is different from that in Shenzhen.

3. The supply level of Shanghai company in May is higher than that in January, while that of Shenzhen Company in May is lower than that in January.

### Table 1: Classification of S&D based on the K-means algorithm.

<table>
<thead>
<tr>
<th>Level</th>
<th>TSH</th>
<th>TSZ</th>
<th>CSH</th>
<th>CSZ</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Feb</td>
<td>Feb</td>
<td>Feb</td>
<td>Feb</td>
</tr>
<tr>
<td>2</td>
<td>Jan, Mar, Apr</td>
<td>Mar, Apr, May</td>
<td>Jan, Mar, Apr</td>
<td>Mar, Apr, May</td>
</tr>
<tr>
<td>3</td>
<td>May, Jun, Jul, Dec</td>
<td>Jan</td>
<td>May, Jun, Jul, Dec</td>
<td>Jan</td>
</tr>
<tr>
<td>5</td>
<td>Nov</td>
<td>Oct, Nov, Dec</td>
<td>Nov</td>
<td>Oct, Nov, Dec</td>
</tr>
</tbody>
</table>

### Table 2: Source data from TopChains used in ARM analysis.

<table>
<thead>
<tr>
<th>Number</th>
<th>Time</th>
<th>TT (yuan)</th>
<th>TSH (yuan)</th>
<th>TSZ (yuan)</th>
<th>TC (yuan)</th>
<th>CSH (yuan)</th>
<th>CSZ (yuan)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1/2013</td>
<td>11458925</td>
<td>5566535</td>
<td>5892390</td>
<td>9458822</td>
<td>4741391</td>
<td>4717431</td>
</tr>
<tr>
<td>2</td>
<td>2/2013</td>
<td>6602000</td>
<td>3542949</td>
<td>3059051</td>
<td>5498476</td>
<td>3115083</td>
<td>2383393</td>
</tr>
<tr>
<td>3</td>
<td>3/2013</td>
<td>10931162</td>
<td>5577172</td>
<td>5353990</td>
<td>8979733</td>
<td>4762488</td>
<td>4217244</td>
</tr>
<tr>
<td>4</td>
<td>4/2013</td>
<td>10656423</td>
<td>5624210</td>
<td>5032213</td>
<td>8519706</td>
<td>4759216</td>
<td>3760490</td>
</tr>
<tr>
<td>5</td>
<td>5/2013</td>
<td>10663865</td>
<td>5236888</td>
<td>5427176</td>
<td>8896461</td>
<td>4706708</td>
<td>4189753</td>
</tr>
<tr>
<td>6</td>
<td>6/2013</td>
<td>11153525</td>
<td>5614077</td>
<td>5539447</td>
<td>9132104</td>
<td>4857657</td>
<td>4274447</td>
</tr>
<tr>
<td>7</td>
<td>7/2013</td>
<td>12761260</td>
<td>5468888</td>
<td>7296372</td>
<td>10337689</td>
<td>4716485</td>
<td>5621205</td>
</tr>
</tbody>
</table>

### Table 3: Discrete variables for Apriori analysis.

<table>
<thead>
<tr>
<th>Number</th>
<th>Month</th>
<th>TT</th>
<th>TSH</th>
<th>TSZ</th>
<th>TC</th>
<th>CSH</th>
<th>CSZ</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Jan</td>
<td>TT6</td>
<td>TSH7</td>
<td>TSZ6</td>
<td>TC5</td>
<td>CSH6</td>
<td>CSZ7</td>
</tr>
<tr>
<td>2</td>
<td>Feb</td>
<td>TT12</td>
<td>TSH12</td>
<td>TSZ12</td>
<td>TC12</td>
<td>CSH12</td>
<td>CSZ12</td>
</tr>
<tr>
<td>3</td>
<td>Mar</td>
<td>TT9</td>
<td>TSH10</td>
<td>TSZ5</td>
<td>TC9</td>
<td>CSH4</td>
<td>CSZ9</td>
</tr>
<tr>
<td>4</td>
<td>Apr</td>
<td>TT11</td>
<td>TSH11</td>
<td>TSZ3</td>
<td>TC11</td>
<td>CSH5</td>
<td>CSZ11</td>
</tr>
<tr>
<td>5</td>
<td>May</td>
<td>TT10</td>
<td>TSH9</td>
<td>TSZ8</td>
<td>TC10</td>
<td>CSH8</td>
<td>CSZ10</td>
</tr>
<tr>
<td>6</td>
<td>Jun</td>
<td>TT7</td>
<td>TSH8</td>
<td>TSZ4</td>
<td>TC7</td>
<td>CSH2</td>
<td>CSZ8</td>
</tr>
<tr>
<td>7</td>
<td>Jul</td>
<td>TT3</td>
<td>TSH4</td>
<td>TSZ7</td>
<td>TC3</td>
<td>CSH7</td>
<td>CSZ4</td>
</tr>
</tbody>
</table>

### Table 4: Descriptions of variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Data type</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_1$</td>
<td>Total customer demand intensity</td>
<td>Total turnover</td>
</tr>
<tr>
<td>$X_2$</td>
<td>Total supply level</td>
<td>Total cost</td>
</tr>
<tr>
<td>$X_3$</td>
<td>Customer demand intensity of Shanghai company</td>
<td>Turnover of Shanghai company</td>
</tr>
<tr>
<td>$X_4$</td>
<td>Supply level of Shanghai company</td>
<td>Cost of Shanghai company</td>
</tr>
<tr>
<td>$X_5$</td>
<td>Customer demand intensity of Shenzhen company</td>
<td>Turnover of Shenzhen company</td>
</tr>
<tr>
<td>$X_6$</td>
<td>Supply level intensity of Shenzhen company</td>
<td>Cost of Shenzhen company</td>
</tr>
</tbody>
</table>

3.2. S&D Level Evaluation Model Based on Apriori. For better presentation and understanding of the level of S&D, Apriori was conducted. Figure 5 reflects the result of monthly S&D
level, which is basically consistent with that of Section 3.1, with more comprehensive presentation and clear comparison of S&D levels among each month.

It should be noted that little research used ARM to deal with logistics expense data since it is marginally difficult to transform the data into the proper form due to the requirements of discretization for data, while this paper made it by ranking the original data and assigning the new values (from 1 to 12) to them; with regard to the results of Apriori, a simple linear model was proposed to calculate the specific value of supply and demand, and the results provided source data for the TOPSIS entropy method.

The main findings can be summarized as follows:

### Table 5: Results of Apriori analysis.

<table>
<thead>
<tr>
<th>Ranking level</th>
<th>Month</th>
<th>TT Confidence level (%)</th>
<th>...</th>
<th>Month</th>
<th>CSZ Confidence level (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Sept, Nov, Dec</td>
<td>33.3</td>
<td>...</td>
<td>Dec</td>
<td>50</td>
</tr>
<tr>
<td>2</td>
<td>Aug, Nov</td>
<td>33.3</td>
<td>...</td>
<td>Jan, Aug, Oct</td>
<td>16.7</td>
</tr>
<tr>
<td>3</td>
<td>July, Oct</td>
<td>33.3</td>
<td>...</td>
<td>Oct, Nov</td>
<td>33.3</td>
</tr>
<tr>
<td>4</td>
<td>Sept, Nov</td>
<td>16.7</td>
<td>...</td>
<td>Jun, Aug</td>
<td>16.7</td>
</tr>
<tr>
<td>5</td>
<td>Oct</td>
<td>50</td>
<td>...</td>
<td>Jul</td>
<td>50</td>
</tr>
<tr>
<td>6</td>
<td>May, Sept, Dec</td>
<td>16.7</td>
<td>...</td>
<td>Mar, Aug, Sept</td>
<td>16.7</td>
</tr>
<tr>
<td>7</td>
<td>Aug</td>
<td>33.3</td>
<td>...</td>
<td>Aug</td>
<td>33.3</td>
</tr>
<tr>
<td>8</td>
<td>Jan, Jul, Oct, Dec</td>
<td>16.7</td>
<td>...</td>
<td>Apr, Sept, Nov, Dec</td>
<td>16.7</td>
</tr>
<tr>
<td>9</td>
<td>Jun</td>
<td>33.3</td>
<td>...</td>
<td>Oct</td>
<td>33.3</td>
</tr>
<tr>
<td>10</td>
<td>Jan, Mar, Jul, Aug</td>
<td>16.7</td>
<td>...</td>
<td>Jan, May, Jun</td>
<td>16.7</td>
</tr>
<tr>
<td>11</td>
<td>Apr</td>
<td>50</td>
<td>...</td>
<td>Apr</td>
<td>33.3</td>
</tr>
<tr>
<td>12</td>
<td>Jan, Mar, Jul</td>
<td>16.7</td>
<td>...</td>
<td>Jun</td>
<td>16.7</td>
</tr>
</tbody>
</table>

### Table 6: Scores of S&D of TopChains.

<table>
<thead>
<tr>
<th>Month</th>
<th>$X_1$</th>
<th>$X_2$</th>
<th>$X_3$</th>
<th>$X_4$</th>
<th>$X_5$</th>
<th>$X_6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan</td>
<td>-2.000</td>
<td>-2.000</td>
<td>-2.500</td>
<td>-2.667</td>
<td>-0.500</td>
<td>-0.700</td>
</tr>
<tr>
<td>Feb</td>
<td>-6.000</td>
<td>-6.000</td>
<td>-5.833</td>
<td>-5.667</td>
<td>-5.833</td>
<td>-6.000</td>
</tr>
<tr>
<td>Apr</td>
<td>-4.000</td>
<td>-4.167</td>
<td>-2.667</td>
<td>-3.000</td>
<td>-3.333</td>
<td>-3.333</td>
</tr>
<tr>
<td>May</td>
<td>-1.500</td>
<td>-1.667</td>
<td>-0.333</td>
<td>-0.167</td>
<td>-2.000</td>
<td>-2.000</td>
</tr>
<tr>
<td>Jun</td>
<td>-1.000</td>
<td>-0.667</td>
<td>-0.833</td>
<td>-0.167</td>
<td>-1.671</td>
<td>-0.333</td>
</tr>
<tr>
<td>Jul</td>
<td>0.667</td>
<td>0.833</td>
<td>-0.333</td>
<td>-0.333</td>
<td>1.167</td>
<td>1.333</td>
</tr>
<tr>
<td>Aug</td>
<td>2.167</td>
<td>2.333</td>
<td>2.000</td>
<td>2.000</td>
<td>1.500</td>
<td>2.833</td>
</tr>
<tr>
<td>Sept</td>
<td>3.833</td>
<td>3.833</td>
<td>2.667</td>
<td>2.833</td>
<td>3.333</td>
<td>3.000</td>
</tr>
<tr>
<td>Oct</td>
<td>2.500</td>
<td>2.500</td>
<td>2.000</td>
<td>1.833</td>
<td>2.500</td>
<td>2.333</td>
</tr>
<tr>
<td>Nov</td>
<td>4.667</td>
<td>4.500</td>
<td>4.833</td>
<td>833</td>
<td>3.500</td>
<td>3.167</td>
</tr>
<tr>
<td>Dec</td>
<td>3.500</td>
<td>3.500</td>
<td>3.667</td>
<td>1.830</td>
<td>2.830</td>
<td>3.167</td>
</tr>
</tbody>
</table>

### Table 7: Scores of three indicators in 12 months.

<table>
<thead>
<tr>
<th>Month</th>
<th>$Y_{p1}$</th>
<th>$Y_{p2}$</th>
<th>$Y_{p3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan</td>
<td>1.000</td>
<td>0.938</td>
<td>0.714</td>
</tr>
<tr>
<td>Feb</td>
<td>1.000</td>
<td>0.971</td>
<td>0.972</td>
</tr>
<tr>
<td>Mar</td>
<td>0.944</td>
<td>1.000</td>
<td>0.889</td>
</tr>
<tr>
<td>Apr</td>
<td>0.960</td>
<td>0.889</td>
<td>1.000</td>
</tr>
<tr>
<td>May</td>
<td>0.900</td>
<td>0.500</td>
<td>1.000</td>
</tr>
<tr>
<td>Jun</td>
<td>0.667</td>
<td>0.200</td>
<td>0.500</td>
</tr>
<tr>
<td>Jul</td>
<td>0.800</td>
<td>1.000</td>
<td>0.875</td>
</tr>
<tr>
<td>Aug</td>
<td>0.929</td>
<td>1.000</td>
<td>0.529</td>
</tr>
<tr>
<td>Sept</td>
<td>1.000</td>
<td>0.941</td>
<td>0.900</td>
</tr>
<tr>
<td>Oct</td>
<td>1.000</td>
<td>0.917</td>
<td>0.933</td>
</tr>
<tr>
<td>Nov</td>
<td>0.964</td>
<td>1.000</td>
<td>0.905</td>
</tr>
<tr>
<td>Dec</td>
<td>1.000</td>
<td>0.499</td>
<td>0.894</td>
</tr>
</tbody>
</table>

### Table 8: Standardization of the entropy weight matrix.

<table>
<thead>
<tr>
<th>Month</th>
<th>$Y_{p1}$</th>
<th>$Y_{p2}$</th>
<th>$Y_{p3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan</td>
<td>1.0000</td>
<td>0.9225</td>
<td>0.4280</td>
</tr>
<tr>
<td>Feb</td>
<td>1.0000</td>
<td>0.9638</td>
<td>0.9440</td>
</tr>
<tr>
<td>Mar</td>
<td>0.8318</td>
<td>1.0000</td>
<td>0.7780</td>
</tr>
<tr>
<td>Apr</td>
<td>0.8799</td>
<td>0.8613</td>
<td>1.0000</td>
</tr>
<tr>
<td>May</td>
<td>0.6997</td>
<td>0.3750</td>
<td>1.0000</td>
</tr>
<tr>
<td>Jun</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Jul</td>
<td>0.3994</td>
<td>1.0000</td>
<td>0.7500</td>
</tr>
<tr>
<td>Aug</td>
<td>0.7868</td>
<td>1.0000</td>
<td>0.0580</td>
</tr>
<tr>
<td>Sept</td>
<td>1.0000</td>
<td>0.9263</td>
<td>0.8000</td>
</tr>
<tr>
<td>Oct</td>
<td>1.0000</td>
<td>0.8963</td>
<td>0.8660</td>
</tr>
<tr>
<td>Nov</td>
<td>0.8919</td>
<td>1.0000</td>
<td>0.8100</td>
</tr>
<tr>
<td>Dec</td>
<td>1.0000</td>
<td>0.3738</td>
<td>0.7880</td>
</tr>
</tbody>
</table>
From February to April, the level of S&D of Top-Chains was not that significant, among which February tended to the worst, exactly consistent with the findings acquired in Section 3.1. Based on these, Top-Chains is recommended to cut down its expenses in these months in a planned way.

From September to December, the level of S&D appeared to be relatively high. The peak demand for customers occurred frequently in November, showing high consistency with the conclusions in Section 3.1, which can provide a prerequisite for the firm to conduct an acceptable increase of investment of human and material resources with the purpose of shortening the time for enterprises to increase the efficiency and quality of customer service, and thus to enhance the customer stickiness. Customer service is significant for 3PL enterprises. Relevant research shows that reducing the loss rate of customers by 5% means that the income of the enterprise can increase by 60%–95% [47].

Compared with the first six months, the level of S&D in the second was higher. The company should consider rationalizing the distribution of the financial resources between the first six months and the second, subsequently reducing the waste of the cost in the vehicle allocation, warehouse leasing and personnel allocation, and improving the operational efficiency.

3.3. Evaluation of S&D Matching Degree Based on the TOPSIS Entropy Weight Method. Figure 6 reflects the evaluation results of S&D matching degree of Top-Chains, with main findings concluded as follows:

(1) From February to April, the level of S&D of Top-Chains was not that significant, among which February tended to the worst, exactly consistent with the findings acquired in Section 3.1. Based on these, Top-Chains is recommended to cut down its expenses in these months in a planned way.

(2) From September to December, the level of S&D appeared to be relatively high. The peak demand for customers occurred frequently in November, showing high consistency with the conclusions in Section 3.1, which can provide a prerequisite for the firm to conduct an acceptable increase of investment of human and material resources with the purpose of shortening the time for enterprises to increase the efficiency and quality of customer service, and thus to enhance the customer stickiness. Customer service is significant for 3PL enterprises. Relevant research shows that reducing the loss rate of customers by 5% means that the income of the enterprise can increase by 60%–95% [47].

(3) Compared with the first six months, the level of S&D in the second was higher. The company should consider rationalizing the distribution of the financial resources between the first six months and the second, subsequently reducing the waste of the cost in the vehicle allocation, warehouse leasing and personnel allocation, and improving the operational efficiency.

(1) The average S&D matching degree of Top-Chains reached 72.27%. The matching in the first quarter was satisfactory (mean = 83.14%), while it had an unsatisfactory decline by 29.5% in the second quarter, and the overall matching in the third and fourth quarters was relatively acceptable (74.4% and 77.89%).

(2) The matching degree in June was extremely undesirable (11.00%). According to the result of Section
3.1, it might be consequent on the oversupply of the subsidiary in Shanghai.

(3) In January, May, June, August, and December, the S&D matching degree was prone to be below the average. In addition, the middle and the end of a year are especially worthy of being paid more attention due to the especially disappointing S&D matching degree, which demonstrated that the efficiency might be unstable, presumably for the deficiency in effective management in these months in terms of external risks, such as fluctuated customer demand and excellent rivals, as well as internal risks.

4. Discussion

There are abundant mathematical methods applied in the field of logistics for the exploration of supply chain and thus to enhance its performance from all aspects such as service, stability, flexibility, and sustainability, e.g., [48–50]. 3PL, as a dominant role in regulating the operation of the supply chain, is also entrusted with great research importance. A few scholars have established the evaluation method for assessing the performance of 3PL enterprises, of which AHP is mostly used, e.g., [51]; [23, 52].

AHP is desirable for evaluation owing to its comprehensive consideration of indicators, e.g., [1, 30], but it simultaneously brings about challenges for data collecting. Also, the weight of index is man-made, which is unavoidably subjective, causing the reliability of the results. For the evaluation of emerging 3PL companies with small or medium size, it is unrealistic due to the high burden to obtain and calculate data. Yu [53] has established a three-stage DEA model for evaluating the efficiency from a financial perspective; however, it is too macroscopic to provide the insight and suggestions for the improvement of the performance of 3PL enterprises. It is necessary to develop a new evaluation method with capacity to process data simply but powerfully, as well as to find the useful information for the improvement of the efficiency, that is why the DM algorithm was applied in the current research.

The main purposes of this study were to establish a feasible quantitative evaluation method to assess the efficiency of emerging nongovernmental 3PL companies measured by the relationship of S&D level, to explore the...
valuable information for the management of 3PL business, and to justify that K-means and ARM could also serve as effective algorithms for the exploration of the efficiency of 3PL in terms of the relationship of the S&D.

By preprocessing the historical data of TopChains with different algorithms, the monthly characteristics of S&D of the company in different time periods and different regions were concluded. Three different algorithms (i.e., K-means, Apriori, and the TOPSIS entropy weight method) were integrated to quantify the monthly S&D level and the matching degree reflected by the turnover and cost of the logistics enterprise through transforming the data into the suitable forms for the data mining algorithm. Compared with the indicators such as regional freight volumes and warehouse capacity, the turnover and cost are more feasible and disposable for nongovernmental 3PL companies, and choosing month as a calculating unit can increase the data availability since the yearly data cannot reflect detail variations and daily data are too temporal, dynamic, and fragile to ensure the reliability of analyses, simultaneously increasing the unnecessary calculation burden.

For the measurement of the efficiency of 3PL companies, the concept of “matching degree of S&D” was proposed as an indicator to determine whether it is balanced between the supply and demand capacity in an emerging 3PL company. Zhang and Wang [45] have defined the matching degree of logistics S&D in the case of expense, time, and quantity based on the logistics network, but this paper understood the notion from another standpoint, i.e., the individual efficiency, related to the capacity of 3PL companies to sustain the balance between supply and demand. Additionally, the results demonstrated that there were fluctuations existing in S&D level of TopChains, for example, both supply and demand levels in February appeared the worst performance. The efficiency of TopChains in terms of supply and demand reached 72.27%, higher than the average value of the industry (66.2%) calculated by Yu [53]. However, it should be noted that due to the differences of the evaluating systems, the comparison might be arguable, which deserves further exploration.

Some interesting insights have also been explored with the DM algorithm, e.g., although with the lowest S&D level, the matching degree in February was the most satisfactory compared with other months, indicating that the value of S&D level might be inconsistent with the matching degree.

The current study is the first to evaluate the efficiency of emerging 3PL companies in terms of supply and demand based on K-means and ARM. K-means is of great popularity owing to its practical, simple, and intuitive characteristics, leading to a broad application in many fields, e.g., market segmentation [54], location determination [55], and document clustering [56]. Also, in the logistics domain, some scholars used K-means to explore the customer relationship management (CRM) [57] and determine the optimum location [58]. However, the direction of K-means applied in the field of logistics is marginally narrow, which is of necessity to be broadened. This paper did make some attempts. The findings of K-means in the paper showed that the S&D level had the propensity to vary with time and space, which can be verified by the results of Apriori. It should be noted that K-means is desirable and effective for the classification of S&D level, while it does have some weaknesses in further analysis. However, ARM can compensate for this deficiency. Little research used ARM to deal with logistics expense data since it is marginally difficult to transform the data into the proper form due to the requirements of discretization for data, while this paper made it by ranking the original data and assigning the corresponding level (from 1 to 12) to them; with regard to the results of Apriori, a simple linear model was proposed to calculate the supply and demand level for the better presentation of the value. Besides, Zhao [59] also integrated Apriori and clustering algorithm in the logistics field to make the strategies of the allocation of goods in the warehouse, with the findings showing that better picking efficiency is acquired. It can be assumed that, in addition to making assessment, ARM and K-means can also serve as powerful tools for the optimization in practical operation.

5. Conclusions and Limitations

The main contributions of the article can be concluded as follows:

(1) A feasible quantitative evaluation method was established to assess the efficiency of emerging nongovernmental 3PL companies. Compared with the traditional evaluation method, constraint by the data type and amount, the evaluation method is more objective, flexible, and effective and has a powerful capability of data processing.

(2) By employing the DM algorithms (i.e., K-means and Apriori), some valuable information on the operation of the business was found. The demand of customers in February has a propensity to decrease substantially, and with the approach of the end of year, the demand shows obvious upward trend while the supply is not that compatible.

(3) The successful application of K-means and Apriori in the current paper confirmed that ARM and K-means could be used as powerful instruments for the analysis of S&D characteristics of 3PL companies, and this attempt broadened the direction of ARM and K-means being applied in the logistics field.

Besides, it deserves noting that the establishment of the advanced information system is of critical necessity for the collection, excavation, and analysis of the history data, which can be utilized to evaluate and thus enhance the efficiency, improve the quality of customer service, and eventually diminish the cost and increase the profits.

There may be some possible limitations of this research, which can be further explored in future research:

(1) In this paper, only one 3PL enterprise was analyzed, and there might be differences between different 3PL companies. Therefore, it is recommended to consider the years of establishment, the number of subsidiaries, and the number of employees in the future research [53].
(2) The evaluation method and required data in the study are different from others, which caused the challenge for the quantitative comparison of the results. Hence, the quantitative comparison with other evaluation methods should be considered for future research.

Data Availability

The data from TopChains used to support the findings of this study have not been made available because of the business risk.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors would like to thank TopChains International Logistics Co., Ltd, for providing the data in the study. The authors would like to thank TopChains International Logistics Co., Ltd, for providing the data in the study.

The authors would like to thank TopChains International Logistics Co., Ltd, for providing the data in the study.

The authors would like to thank TopChains International Logistics Co., Ltd, for providing the data in the study.

References


