An Improved K-means Algorithm for Supplier Evaluation and Recommendation of Purchase and Supply Platform

Junting Li 1,a, Wenwen Qiu 1,b, Weigang Li 2,c

1 Xi’an Shiyou University School of Economics and Management Xi’an, Shanxi, China
2 Northwestern Polytechnical University School of Software Xi’an, Shanxi, China

a jtinglee@xsyu.edu.cn, b 1720568688@qq.com, c liweigang@nwpu.edu.cn

Abstract. Aiming at the problem of supplier evaluation and selection in B2B e-commerce, a supplier evaluation and recommendation method based on improved k-means algorithm is proposed. Firstly, this paper analyzes the supplier evaluation and recommendation ideas based on the purchase and supply platform, and proposes the data mining algorithm ideas of clustering analysis and AHP evaluation; secondly, K-means algorithm is proposed based on the data mining model, and the algorithm is optimized according to the data characteristics of the purchase and supply platform; finally, taking the business data of the industrial product purchase platform of the volume purchase network as an example, not only the effectiveness of the algorithm is verified, but also the clustering effect of the algorithm is good and the calculation speed is fast, which provides a practical and effective supplier evaluation and recommendation method for B2B trading website.

Keywords: E-commerce; Purchase and supply platform; K-means algorithm; Supplier evaluation and recommendation.

1. Introduction
With the promotion of economic globalization, in the increasingly competitive market today, the game between enterprises is not only the competition of enterprises’ own strength, but also directly related to the suppliers who cooperate with enterprises directly. Especially in the 21st century, the rapid development of the Internet makes more and more purchasing enterprises rely on convenient network resources and use some B2B e-commerce platforms to purchase. However, B2B online trading platform suppliers are mixed, trading information asymmetry, it is difficult to ensure that buyers can choose the right supplier every time, this requires the platform itself to manage and evaluate suppliers.

Along with the change of business model and the development of integrated supply chain management theory, as the supplier evaluation and selection problem of the classic problem of resource optimization allocation, it is constantly endowed with new connotation——supplier evaluation and recommendation of data mining of historical trading platform. At the same time, it has been paid more and more attention by academic and business circles. Now supplier evaluation has adopted different evaluation methods for different situations, such as the new hesitant multi-attribute decision model of supplier selection based on HTFWGBM operator [1]; in order to alleviate expert experience to evaluate
supplier performance, an intelligent supplier evaluation model based on data-driven support vector regression (SVR) is proposed [2]; for the optimal supplier selection in fuzzy environment, based on the intuitionistic fuzzy set hyperlink induced topic search method [3], for the supplier selection decision-making of subjective and objective mixed evaluation information, the group decision-making method based on intuitionistic fuzzy cross entropy and grey correlation [4]. These evaluation methods basically focus on the optimization of methods, but there are massive and dynamic supplier data in B2B trading platform, how to quickly and conveniently screen, classify and evaluate them is an important issue to be solved in supplier selection. A large number of scholars such as Merve Er Kara [5] and Park [6] began to apply data mining methods and genetic algorithms to supplier evaluation, deeply mining the hidden relationship between massive data in B2B platform database, and finally obtained the ordering rules of suppliers, which provided a new research perspective for supplier selection.

2. Supplier evaluation indicators and data mining models for the platform

2.1. Supplier Cluster Analysis and Evaluation Indicators
The supplier evaluation of the platform can be based on the index of transaction history data, and its evaluation is more objective and comprehensive, which can reflect the favorable conditions of each supplier. Establishing the index of supplier selection and evaluation can not only use cluster analysis as the basis for dividing cluster classes, but also form an index system for further evaluation to complete the cluster analysis and effective evaluation of suppliers. The supplier evaluation index involves the factors of enterprise economic status, quality management, technical level and so on. Dickson [8] was the first American scholar to carry out systematic research on it, and put forward 23 supplier evaluation criteria with quality, delivery time and historical performance as the top three important factors. At the same time, Ma Shihua [9] from the perspective of grey system theory, discusses the shortcomings of the existing evaluation methods to determine the index weight, and finally concludes that the top three factors of transaction are quality, cost and delivery, which are consistent with the general research results. And for the supplier evaluation on the e-commerce platform, Wang Xuping [10] proposed that business capability, informatization degree, information technology and service level of suppliers should be taken as evaluation criteria. Therefore, based on the research results of the existing supplier evaluation system, combined with the general business process of the procurement and supply platform and the business data in the database, a three-level comprehensive evaluation index system is constructed, as shown in Figure 1.

![Image of Supplier evaluation index]

Fig.1 Supplier evaluation index
2.2. Supplier Selection Evaluation and Data Mining by Purchasers

When purchasing suppliers in the platform, because the types of items purchased are different, the types of suppliers participating in the bidding will be different, supplier evaluation also needs to be carried out for different types of suppliers. At the same time, it is difficult for buyers to determine the level of suppliers participating in each bid in advance, when the platform recommends suppliers to buyers, it can only select the best suppliers from this batch of suppliers. Therefore, this classification of suppliers can not be "guided learning" classification method.

In the supply platform, due to the long trading behavior of buyers and different suppliers, there is a large number of basic information of suppliers and related information of supplier transactions in the database, there is a certain relationship between these data, and it can also reflect the information of the supplier in terms of the price, quality and logistics time. However, it is difficult for buyers to obtain the detailed transaction data of the trading platform to evaluate suppliers, and the managers of the trading platform can not directly evaluate the suppliers. Because a large number of transaction data are distributed in different tables of the database, it is necessary to integrate and analyze these data. Therefore, the method of data mining can be used to discover hidden knowledge. The clustering algorithm in data mining is a classification method without knowing the result in advance, it can classify according to the characteristics of the current data, and the purchasing and supplying platform selects the best supplier from the suppliers who have completed the classification.

2.3. Algorithm for Data Mining

Generally, the data mining of the platform is related to the system or specific business, so it is specific and not universal. In order to unify the standard of data mining and guide the process of data mining, the general data mining model shown in Figure 2 reveals the logical relationship of data mining manager, data preprocessing, mining kernel and so on.

![Fig.2 General data mining model](image)
Based on the data mining model of figure 2, the mining algorithm can be processed as an independent module [7]. The algorithm of supplier evaluation and recommendation can be divided into two steps: first, the method of cluster analysis of suppliers participating in bidding quotation is classified, and suppliers are divided into several different classes according to evaluation index, then the weight value of the index is calculated by analytic hierarchy process, and the optimal supplier is found according to the weighted score of the index. The specific analysis process is as follows:

1. Cluster analysis

Using clustering analysis algorithm to analyze the suppliers involved in bidding, the suppliers are divided into different clusters, so that the similarity of each index data of suppliers in the same cluster is high, and the difference of each index data of suppliers between clusters and clusters is high. The index comes from the established supplier evaluation system, and selects several attributes for cluster analysis from the evaluation system, and divides the suppliers into different categories according to the attributes.

2. Analytic hierarchy process evaluation

After cluster analysis of the selected suppliers, the weight of the first level index relative to the target layer is calculated by AHP, then the relative weight of the second level index to the first level index is calculated, and then the absolute weight of the second level index is calculated. The score of each category is calculated according to the weight and index value, and the optimal supplier group is selected. Finally, the score of each supplier is calculated one by one from the optimal supplier group, and the optimal supplier is selected to recommend to the purchaser.

3. Supplier cluster analysis algorithm based on improved K-means

Based on the above analysis of suppliers of procurement and supply platform and put forward the idea of supplier evaluation algorithm based on data mining, an improved K-means algorithm is proposed.

3.1. Description of Algorithm Ideas and Steps

K-means algorithm, also known as the K- average algorithm, is an important implementation method in clustering mining, its basic idea is partitioning methods [11]. For data sets containing n objects, find out k cluster centers and divide the data in the data set into k clusters according to the principle of proximity, so that the k clusters meet two requirements: the number of data in each cluster must not be less than one, each data object belongs to and belongs to only one cluster, and then adjust the data attribution according to the decision function [12].

The algorithm step is to select k objects randomly from the n data objects to be classified first, the k objects represent the average value in the cluster at the time of the first classification, the euclidean distance between the remaining n-k objects and the k objects is calculated, the data object with the smallest distance and the mean object are divided into the same cluster, so that n data are divided into k different clusters. Then, the average value in each cluster is recalculated, and the above classification steps are iterated until the result of the last classification is the same as that of the penultimate classification, and the algorithm is convergence of the criterion function [13]. The convergence function is shown in the following equation (1):

$$E = \sum_{i=1}^{k} \sum_{x \in c_i} |x - \bar{x}_i|^2$$

(1)

Where E represents the sum of squares difference of all data objects, $x$ represents data objects, $\bar{x}_i$ represents the mean value within the cluster $c_i$. The smaller the convergence function, the more similar the data in the cluster, the higher the independence of the cluster and the lower the similarity correlation between the cluster and the cluster. The euclidean distance [14] is used to calculate the distance of the data object to the center point of each cluster. The distance formula is as follows (2):

$$d(x,y) = \left[ \sum_{i=1}^{n} |x_i - y_i|^2 \right]^{1/2}$$

(2)
Where, \( x, y \) is the corresponding feature, \( n \) is the dimension of the feature.

Describe the steps of the K-means algorithm below.

**Algorithm:** \( K-means(S, K) \)

**Input:** \( k \) number of sets \( S \), target categories containing \( k \) data objects

**Output:** \( k \) clusters

**Steps:**

1. \( k \) data objects are randomly selected from the set \( S \) as the average value \( P_1, ..., P_k \) of the initial \( k \) clusters \( C_1, ..., C_k \)
2. Iterate the following steps for each data object in the data collection \( S \)
   - Calculate the distance from the data object to the mean \( P_i \) of each cluster and find the nearest \( P_i \) to the data object \( O \)
   
   \[
   \|O - P_j\| = \min \{\|O - P_i\| : j \leq k\}
   \]

   Assign data object \( O \) to the cluster \( C_i \) to which \( P_i \) belong
   - For each cluster \( C_j \) (\( 1 \leq j \leq k \))
   
   \[
   P_j = M_j = \frac{\sum_{i \in C_j} X_i}{\sum_{i \in C_j}}
   \]

   // Calculate the average of objects within each cluster

   \[
   E = \sum_{i=1}^{k} \sum_{o \in C_i} (x - M_i)^2 / \sum_{i=1}^{k} \sum_{o \in C_i}
   \]

3. Until \( k \) clusters no longer change, that is, the criterion function \( E \) converges

**3.2. K-means improvement algorithm**

K-means algorithm has the advantages of fast and efficient, good clustering effect and suitable for the processing of large data sets, but there are also some defects, such as the difficulty of selecting the initial clustering center [14]; the same degree of view of each variable in the clustering process, and frequent database traversal; sensitive to noise data [15] and so on. Aiming at this deficiency in the K-means algorithm, the following improvements are made:

1. Create a two-dimensional array \( \text{sum}[k][m] \) to hold data objects within each cluster and a one-dimensional array \( \text{count}[k] \) to hold the number of data objects within each cluster. By this method, the new mean value of each cluster can be calculated directly by formula
   
   \[
   \sum_{j=0}^{m} \text{sum}[i][j] / \text{count}[i], 0 < i \leq k
   \]

   without traversing the database.

2. The condition at the end of the K-means algorithm is modified to mean that the average value of each cluster is no longer changed or significantly changed, so that the average value of clusters can be calculated using two-dimensional array \( \text{sum}[k][m] \) and one-dimensional array \( \text{count}[k] \) instead of traversing the database calculation criterion function.

   Based on the above improved design, the following improved algorithm is given:

   **Algorithm:** improvement \( k-means(S,k) \)

   **Input:** the set \( S \) with \( n \) data objects and the target classification number \( k \)

   **Output:** \( k \) clusters, the average of each cluster no longer changes

   **Steps:**

   1. \( k \) data objects are randomly selected from the set \( S \) as the average value \( P_1, ..., P_k \) of the initial \( k \) clusters \( C_1, ..., C_k \)

   2. Initialize the sum[i][j]=0;count[i]=0; // sum[i][j] is used to save the sum of the j attribute value of the data object in the i cluster (\( 0 \leq i \leq k, 0 \leq m \) object dimension) // count[i] is used to record the number of data objects in the i cluster;

   3. Iterate the following steps

      - Cover the old mean with the new mean
      - for \( t=1 \) to \( n \) do 

int j; // save class tags
for i=1 to k {calculate the distance dit (i=1,...,k) between each data object and the mean of each
cluster ;find the smallest d and write down cluster number p }
Assign objects to Cp
\[
\text{Sum} [m][i] = \text{Sum} [m][i] + O[i](i)
\]
(0\leq i<\text{object dimension})
\[
\text{Count} [m] = \text{Count} [m] + 1 ; // the total number of objects of this class accumulates
\]
//for end of cycle
\[
P_m = \frac{\text{Sum} [m][i]}{\text{Count} [m]} \quad (0\leq i<\text{object dimension}); //get the new mean of each cluster
\]
//get the new mean of each cluster
Compare the new and the old;
(4) Until old and new means no longer change.

4. Examples of supplier evaluation and recommendation algorithm

4.1. Experimental Algorithm
In this paper, based on the one-time purchase of the purchasing and supply platform of the quantity
purchasing network, the dynamic initial selection of 10 suppliers is taken as the sample data, and the
improved K-means algorithm is applied to cluster analysis. On the basis of the classification of suppliers,
AHP is used to calculate and score the classified suppliers according to the evaluation index, finally, the
classification and ranking suppliers are recommended to the purchasers according to the scoring results.
The supplier's initial index data is from the quantity acquisition network, and the five indicators of
product quality, product price, product delivery, company reputation and company strength are used for
cluster analysis. The index data is shown in Table 1.

| Supplier  | Product quality | Product price | Product delivery | Corporate reputation | Corporate strength |
|----------|-----------------|---------------|------------------|----------------------|-------------------|
| Supplier A | 8.3             | 8.6           | 7.9              | 7.7                  | 7.9               |
| Supplier B | 7.5             | 7.7           | 6.9              | 6.9                  | 7.3               |
| Supplier C | 6.9             | 7.6           | 6.9              | 6.3                  | 6.8               |
| Supplier D | 7.4             | 7.4           | 6.9              | 6.6                  | 6.9               |
| Supplier E | 7.1             | 6.8           | 8.1              | 7.9                  | 7.3               |
| Supplier F | 7.3             | 7.4           | 7.3              | 7.0                  | 6.9               |
| Supplier G | 7.9             | 7.4           | 6.8              | 6.8                  | 6.7               |
| Supplier H | 7.9             | 7.0           | 7.2              | 6.9                  | 6.9               |
| Supplier I | 8.7             | 8.6           | 8.1              | 8.3                  | 8.4               |
| Supplier J | 7.1             | 7.4           | 6.9              | 6.4                  | 6.3               |

Clustering attributes are set as follows: the value range of initial parameter k is limited to [3, 6]; the
five indexes of product quality grade, product transaction price, delivery time, number of certifications,
and company registered capital are used for cluster analysis; the euclidean distance is used as the
calculation method of distance measurement, and the dissimilarity of data objects is used to cluster.
Cluster analysis was carried out on the data of sample suppliers, the initial parameter k of clustering
algorithm is set as 4, which means the suppliers to be selected are divided into 4 categories.
After cluster analysis, the supplier is assigned to different categories, table 2 shows the variation of
the mean of each class after iteration.
Table 2. Cluster analysis process

| Supplier | Product quality | Product price | Product delivery | Corporate reputation | Corporate strength | Number of suppliers | Number of clusters | Type |
|----------|----------------|--------------|-----------------|----------------------|-------------------|--------------------|------------------|------|
| Supplier A | 0             | 0            | 0               | 0                    | 0                 | 0                  | 0                | 1    |
| Supplier B | 0             | 0            | 0               | 0                    | 0                 | 0                  | 0                | II   |
| Supplier C | 0             | 0            | 0               | 0                    | 0                 | 0                  | 0                | III  |
| Supplier D | 0             | 0            | 0               | 0                    | 0                 | 0                  | 0                | IV   |
| Supplier E | 8.5           | 8.6          | 8.0             | 8                    | 8.2               | 2                  | 1                | I    |
| Supplier F | 7.5           | 7.7          | 6.9             | 6.9                  | 7.3               | 1                  | 1                | II   |
| Supplier G | 7.52          | 7.2          | 7.26            | 7.0                  | 6.9               | 5                  | 1                | III  |
| Supplier H | 7.0           | 7.5          | 6.9             | 6.4                  | 6.6               | 2                  | 1                | IV   |
| Supplier I | 8.5           | 8.6          | 8.0             | 8.0                  | 8.2               | 2                  | 2                | I    |
| Supplier J | 7.5           | 7.7          | 6.9             | 6.9                  | 7.3               | 1                  | 2                | II   |
| Supplier A | 7.52          | 7.2          | 7.3             | 7.0                  | 6.9               | 5                  | 2                | III  |
| Supplier B | 7.0           | 7.5          | 6.9             | 6.4                  | 6.6               | 2                  | 2                | IV   |

It can be seen from Table 2 that after two clustering analysis, the mean value of each index of supplier evaluation has no longer changed. According to the condition of cycle end in the improved clustering algorithm, when the new and old mean values do not change, the end of clustering can be judged. Through the classification of clustering algorithm, the selected suppliers are divided into four categories, in which class I contains 2 suppliers, class II contains 1 supplier, class III contains 5 suppliers and class IV contains 2 suppliers.

Using analytic hierarchy process to evaluate suppliers, the supplier evaluation system is divided into three layers: target layer, first level index layer and second level index layer, in which the first level index layer includes quality, cost, delivery, company reputation, company strength. By establishing the judgment matrix of pairwise relative importance of the first level index, the maximum eigenvalue and eigenvector of the matrix are calculated, and the weight of the first level index relative to the target layer is obtained (the analytic hierarchy process is slightly), as shown in Table 3 below.

Table 3. I indicator weights

| Level I indicators | Product quality | Product price | Product delivery | Corporate reputation | Corporate strength |
|-------------------|-----------------|--------------|-----------------|----------------------|-------------------|
| Weight            | 0.512           | 0.246        | 0.105           | 0.059                | 0.078             |

After calculating the weight of the first level index, we also construct the judgment matrix of the second level index, and get the weight of the second level index relative to the first level index, as shown in Table 4 below.

Table 4. Relative weights of the secondary indicators

| Level II indicators $i_a$ | Product quality | Product price | Product delivery | Corporate reputation | Corporate strength |
|---------------------------|-----------------|--------------|-----------------|----------------------|-------------------|
|                           | 0.58            | 0.84         | 0.6             | 0.63                 | 0.42              |
| Level II indicators $i_b$ | 0.42            | 0.16         | 0.4             | 0.37                 | 0.26              |
| Level II indicators $i_c$ | /               | /            | /               | /                    | 0.32              |

After calculating the relative weight of the secondary index, it is necessary to test the consistency. Through the test, the consistency is 0.09, less than 0.11, which meets the requirement of consistency. Then the absolute weight of the secondary index relative to the target layer is calculated in combination with Tables 3 and 4, and the results are shown in Table 5.
Table 5. Absolute weight of secondary indicators

| Level II indicators $i_a$ | Product quality | Product price | Product delivery | Corporate reputation | Corporate strength |
|---------------------------|-----------------|---------------|------------------|----------------------|-------------------|
| 0.29696                   | 0.20664         | 0.063         | 0.03717          | 0.03276              |
| Level II indicators $i_b$ | 0.21504         | 0.03936       | 0.042            | 0.02183              | 0.02028 |
| Level II indicators $i_c$ | /               | /             | /                | /                    | 0.02496 |

According to the weight of the secondary index, the comprehensive scores of all kinds after cluster analysis are calculated. The calculation formula is: class score = quality score × quality absolute weight + price score × price absolute weight + delivery time score × delivery time absolute weight + certification number score × absolute weight of authentication number + registered fund score × absolute weight of registered funds. The results obtained by calculation are shown in Table 6.

Table 6. Cluster score

| Product quality | Product price | Product delivery | Corporate reputation | Corporate strength | Cluster types | Number of suppliers | Score    |
|-----------------|---------------|------------------|----------------------|-------------------|---------------|---------------------|----------|
| 8.5             | 8.6           | 8                | 8                    | 8.15              | I             | 2                   | 5.369618 |
| 7.5             | 7.7           | 6.9              | 6.9                  | 7.3               | II            | 1                   | 4.748649 |
| 7.52            | 7.2           | 7.26             | 7.04                 | 6.94              | III           | 5                   | 4.6673584 |
| 7               | 7.5           | 6.9              | 6.35                 | 6.55              | IV            | 2                   | 4.5138275 |

Class I scores 5.369618, Class II scores 4.748649, Class III scores 4.6673584, Class IV scores 4.5138275. It is concluded that class I should be the optimal supplier group, a supplier should be selected from I to recommend to the purchasing enterprise. Generally speaking, according to the results of cluster analysis, the purchasing enterprise can choose the suitable supplier according to the actual situation, but the optimal recommendation from the intelligent service of the purchasing platform can be further calculated: class I contains the supplier A and the supplier I, through the two-level index weighted calculation, the final weighted index value of the supplier A is 8.35478, the weighted index value of the supplier I is 8.27861, and supplier A is the best recommended supplier.

5. Conclusion

This paper proposes a supplier evaluation and recommendation method based on the perspective of supply platform, which is different from the supplier selection and evaluation method of general buyer's perspective. The basic data of this method comes from the purchase and supply platform, which has a large number of dynamic and real first-hand transactions, for general purchasers, the data is not easy to obtain, which may form a trend of supplier evaluation in the future. The improved k-means algorithm is used in the research, which improves the shortage of frequent traversing database in traditional algorithm; through the establishment of index system, the cluster center is easy to set, which is not affected by noise data and has practical significance; the simple array definition is used to improve the convergence speed of the algorithm, and the algorithm is simple and practical. The algorithm not only proves that the algorithm has fast calculation speed and good clustering effect, but also reduces the time complexity of the algorithm and improves the solution efficiency through the actual data of the quantity acquisition network; at the same time, it can be applied to practical application, providing a practical and effective algorithm for B2B trading website.
References

[1] Zhang Jian. Study on Supplier Selection Model Based on HTFWGBM Operator [J]. China Management, 2019, 27(03):137-143.

[2] Yijun Cheng, Jun Peng, Xin Gu, Xiaoyong Zhang, Weirong Liu, Zhuofu Zhou, Yingze Yang, Zhiwu Huang. An Intelligent Supplier Evaluation Model Based on Data-driven Support Vector Regression in Global Supply Chain [J]. Computers & Industrial Engineering:1-36[2019-04-26].https://doi.org/10.1016/j.cie.2019.04.047.

[3] Zhijia Yan, Wenting Yang, Xiaolong Huang, Xiangrong Shi, Wenyu Zhang, Shuai Zhang. A Reputation-Enhanced Hybrid Approach for Supplier Selection with Intuitionistic Fuzzy Evaluation Information [J]. Mathematics, 2019, 07(03):1-17.

[4] Han Erdong, Xu Guodong. Decision-making of mixed evaluation information supplier selection based on intuitionistic fuzzy cross-entropy and grey correlation. Science. [J].Technology and Engineering 17(07):1-9.

[5] Er Kara, Seniye Ümit Oktay Firat, Abhijeet Ghadge. A data mining-based framework for supply chain risk management [J]. Computers & Industrial Engineering:1-12[2018-12-17].https://doi.org/10.1016/j.cie.2018.12.017.

[6] Park, Sung Chul; Lee, Jang Hee. Supplier selection and stepwise benchmarking: a new hybrid model using DEA and AHP based on cluster analysis[J]. Journal of the Operational Research Society, 2018, 69(03):449-466.

[7] Zeng Minghua, Wang Yinsong, Yang Xiaoguang. A Study on Evaluation Index System of Supplier Synergy [J]. Computer Engineering and Applications, 2019, 07(03):1-17.

[8] Park, Sung Chul; Lee, Jang Hee. Supplier selection and stepwise benchmarking: a new hybrid model using DEA and AHP based on cluster analysis[J]. Journal of the Operational Research Society, 2018, 69(03):449-466.

[9] Ma Shihua, Wang Xubin. A method for determining the weight of supplier evaluation indicators[J]. Industrial Engineering and Management, 2002(06):5-8.

[10] Wang Xuping, Chen ao. E-commerce-based supplier evaluation and optimization [J]. Management Science, 2004(04):49-53.

[11] Ma Yunhong, Wang Chenghan, Jiang Tengjiao, Zhang Kun. An Automatic Clustering Algorithm Based on Data Inclusion Journal of Northwest Polytechnic University, 2016, 34(05):863-866.

[12] Feng Bo, Hao Wenning, Chen Gang, Zhan Donghui. Optimization of initial cluster center selection of K-means algorithm [J]. Computer Engineering and Applications, 2013, 49(14):182-185 192.

[13] Jiang Shengyi. An enhanced K-means clustering algorithm [J]. Computer Engineering and Science, 2006,28(11): 56-59.

[14] Tang Dongkai, Wang Hongmei, Hu Ming, Liu Gang. Modified K-means Algorithm for Optimizing Initial Cluster Center [J]. Microcomputer Systems, 2018, 39(08):1819-1823.

[15] Shen Yanyun. Classification of B2B Supplier Credibility Based on Data Mining [D]. Dalian Maritime University, 2013.