Research on supply chain planning based on genetic algorithm and long short-term memory

Xu Wang[1], *Yujie* Li[1,*], *Qizong* Lu[1], and *Yuchen* Qiu[2]

[1]School of Artificial Intelligence, Guilin University of Electronic Technology, Guilin, Guangxi, 541000, China
[2]School of Information and Communication, Guilin University of Electronic Technology, Guilin, Guangxi, 541000, China

Abstract. With the integration of intelligent algorithm into the supply chain process, the fficiency of supply chain planning has been further improved through automatic prediction and decision-making. Although intelligent algorithms are developing, their challenges including real-time nature of supply chain planning and the complexity of scenarios hinder their true potential. In this study, we proposed an improved genetic algorithm (GA)-long short-term memory (LSTM) neural network prediction algorithm to solve various optimization planning problems for the supply chain from suppliers to production enterprises. Specifically, to determine stable suppliers, we first constructed the technique for order preference by similarity to ideal solution (TOPSIS) model to quantitatively evaluate each supplier, and the rationality of the index weight of the TOPSIS algorithm can be enhanced by the entropy method. Finally, the GA and LSTM were used to solve the decision-making and planning problem in raw material supply chain. Our results indicate that the algorithm we proposed can not only efficiently solve the decision planning problem in the raw material supply chain, but it also reasonably analyzes the suppliers quantitatively.

Keywords: Data mining, Supply chain planning, Entropy-TOPSIS, GA, LSTM.

1 Introduction

The concept of artificial intelligence is broad and it involves many technologies such as computational, distributed, and cluster intelligence. For the current devices, links that can sense their surroundings and perform tasks based on the principle of maximizing the likelihood of goal completion are being integrated with artificial intelligence. Additionally, artificial intelligence is crucial in supply chain, which is the process of providing goods and services [1]. Specifically, it can promote the network chain structure formed by upstream and downstream enterprises involved in the production and distribution process to provide goods or services to end users [2]. However, it is extremely challenging to achieve efficient

* Corresponding author. Email: yujieli@guet.edu.cn

supply chain management under the complex and changing market environment during economic globalization [3]. Therefore, it is necessary to find an efficient supply chain planning [4]. Since the outbreak of the new crown epidemic in 2019, a fast and accurate prediction of the supply chain can reduce personnel contact and improve service efficiency. Internet of Things (IoT) and its impact on supply chain management (SCM) has grown over time [5], which results to the growth of input data in demand planning methodology. Here, intelligent algorithms have an advantage in data processing and application. It can capture historical shipment data and transform it into forecasts. However, forecasting based on demand only hinders its practical application. That is, intelligent algorithm-based forecasting engines are always looking for the best approach to combine algorithms and data streams for different levels of forecasting problems. Therefore, the combination of artificial intelligence algorithms and supply chain management is crucial for supply chains to improve operational efficiency and enhance sustainability [6]. The main contributions of this study include 3 aspects:

We propose a quantitative supplier evaluation framework based on the entropy-technique for order preference by similarity to ideal solution (TOPSIS) method to improve the reliability of the quantitative evaluation method by constructing more representative features;

By improving the genetic algorithm (GA), the objective planning problem can be solved more efficiently and the optimal ordering scheme can be solved by combining it with long short-term memory (LSTM);

The results of the case study indicate that our improved algorithm is more reliable in quantitative evaluation of suppliers and it is more efficient in calculating the optimal ordering solution.

2 Entropy-TOPSIS quantitative evaluation method

Supplier quantitative evaluation can play a crucial role in supply chain planning. An effective quantitative evaluation method can not only accurately assess the supplier priorities, but it can also improve the model performance significantly. In this study, we proposed a quantitative evaluation framework for suppliers based on Entropy-TOPSIS. Here, we mainly introduce the application of entropy-TOPSIS method in the quantitative evaluation of raw material suppliers, and we discuss the approach for improving the performance of the algorithm.

2.1 Feature indexs

A single feature is often insufficient for obtaining superior quantitative results. It is essential to select superior data features from the original data to obtain accurate quantitative results. Therefore, we processed the original data and constructed representative features such as supply times, satisfaction rate, and variance of satisfaction rate.

Supply times(x_{i1}) of the supplier in time t are expressed as follows:

$$x_{i1} = \sum_{j=1}^{T} \{t = 1 \mid g_{ij} > 1\}, \tag{1}$$

Where g_{ij} is the supply of the $i-th$ supplier in week j.

Total supply of suppliers (x_{i2}) can reflect the supply capacity of the supplier. The total supply volume of the supplier in time t is defined as $x_{i2} = \sum_{j=1}^{T} g_{ij}$.

Variance of supply quantity (x_{i3}) is defined as follows:

$$x_{i3} = \frac{\sum (g_{ij} - \overline{g}_i)^2}{T-1}, \overline{g}_i = \frac{1}{n}\sum_{i=1}^{n} g_{ij}, \tag{2}$$

where $\overline{g}_i$ is the average of supplier i supply. The supply quantity becomes more stable as the variance decreases.

Satisfaction rate (x_{i4}) represents the ratio of the number of deliveries greater than the order quantity to the total number of deliveries. It is defined as follows:

$$x_{i4} = x_1^{-1} \sum\nolimits_{j=1}^{T} \{t = 1 \mid \frac{g_{ij}}{q_{ij}} > 1\}, \tag{3}$$

where q_{ij} is the order quantity of the enterprise to supplier i in week j.

Variance of satisfaction rate (x_{i5}) is calculated. The supplier's supply situation is more stable if the variance of the satisfaction rate is small. Additionally, it can complete the requirements on time and quantity. The variance of satisfaction rate is defined as follows:

$$x_{i5} = \frac{\sum (x_{i4} - \overline{x}_4)^2}{T-1}, \overline{x}_4 = \frac{1}{n}\sum_{i=1}^{n} x_{i4}, \tag{4}$$

where $\overline{x}_4$ is the average of the satisfaction rate of all suppliers.

2.2. Quantitative evaluation of the entropy-TOPSIS method

TOPSIS was developed by C.L. Hwang and K. Yoon [7]. It ranks according to the proximity of a limited number of evaluation objects to the idealized goal. It evaluates the relative advantages and disadvantages of the existing objects. As a sorting method approaching the ideal solution, it only requires that each utility function has a monotonic increasing (or decreasing) property. It is a commonly used and effective method in multi-objective decision analysis, and it is also known as the advantages and disadvantages solution distance method [8]. The quantitative evaluation includes the following steps:

Indicator attribute isotropization: If the data in one dimension is larger compared to the superior and the data in another dimension is smaller compared to the superior, it will cause confusion in scale. Therefore, it is necessary to identify the extremely small indicators in the characteristics, convers cost-based indicators to benefit-based indexes, and the formula is defined as $x' = \frac{1}{x}(x > 0)$. where x is the feature index of the supplier. Similarly, the variance of the satisfaction rate and the variance of the supply are both extremely small indexes. However, others are extremely large indexes and they do not need to be processed.

Constructing normalized initial matrix: Assuming that the number of supplier individuals is m and each individual has n features, the original data matrix is defined as:

Thereafter, we can obtain the normalized matrix, A, by constructing the weighted normalization matrix and normalize the feature vectors.

Determining the optimal and worst solutions: Optimal solution $A+$: consists of the maximum value of each element in each column of A.

The worst solution $A-$: consists of the smallest value of each element in each column of A.

$$X=\begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{mn} \end{pmatrix},\ a_{ij}=x_{ij}\left(\sum_{i=1}^{n}x_{ij}^{2}\right)^{-\frac{1}{2}},\ A=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, \tag{5}$$

$$A^{+}=(\max\{a_{11},\cdots,a_{m1}\},\cdots,\max\{a_{1n},\cdots,a_{mn}\})=A_{1}^{+}+\cdots+A_{n}^{+}, \tag{6}$$

$$A^{-}=(\min\{a_{11},\cdots,a_{m1}\},\cdots,\min\{a_{1n},\cdots,a_{mn}\})=A_{1}^{-}+\cdots+A_{n}^{-}, \tag{7}$$

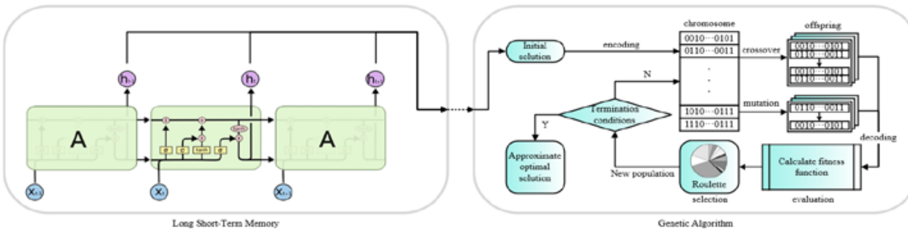


Fig. 1. Model framework.

Weight coefficient determined by the entropy weight method: Entropy weighting method is an objective empowerment method, which can effectively avoid the bias caused by artificial objective factors and obtain the weight of each feature index with high accuracy(Huang 2021). First, the original data matrix is normalized by column for ratio as $c_{ij}=\frac{x_{ij}}{\sum_{i=1}^{m}x_{ij}}$.

Calculate the entropy value of each index: $u_j=-b\sum_{i=1}^{m}c_{ij}\ln c_{ij},(j=1,2,\cdots,m)$, $b=\frac{1}{\ln m}$, where coefficient b is related to the sample size. Therefore, the weighting coefficients of each feature index can be calculated as follows:

$$s_j=\frac{1-u_j}{\sum_{i=1}^{m}(1-u_k)},(j=1,2,\cdots,n), \tag{8}$$

Here, the larger the entropy weight coefficient s_j, the greater the influence represented by the feature index and its effect on the comprehensive evaluation is also greater.

Calculate the distance of each supplier from the optimal solution and the worst solution: The distance of the optimal solution and the worst solution is defined as follows:

$$D_i^{+}=\sum_{j=1}^{n}s_{\mathrm{j}}^{\frac{1}{2}}(A^{+}-a_{ij}),\ \ D_i^{-}=\sum_{j=1}^{n}s_{\mathrm{j}}^{\frac{1}{2}}(A^{-}-a_{ij}), \tag{9}$$

This allows us to calculate the closeness of the supplier T_i to the optimal solution. Where T_i ranges from 0 to $1(0\le T_i\le 1)$, and when it closer to 1 means the better the supplier.

$$T_i=\frac{D_i^{-}}{D_i^{+}+D_i^{-}}, \tag{10}$$

3 Supplier planning model based on GA and LSTM

By quantitatively analyzing suppliers, we can select those that are worthy of a long-term relationship. However, to obtain the maximum commodity profit, we also need to find the optimal ordering strategy; therefore, we solved the optimal supplier planning scheme for the enterprise by combining the LSTM model [10] and GA [11].

3.1 Model constructed based on LSTM

When dealing with events that have a relationship with time, such as time series data including weather conditions on consecutive days, where there is some connection between the weather conditions of the current and the past days, the use of traditional neural networks is often powerless; thus, S Hochreiter et al. proposed LSTM. LSTM is not only effective in dealing with long-term dependence problems, but it can also avoid the gradient vanishing problem of neural networks and speed up the model training. In general, the supplier's supply shows a certain periodic pattern; therefore, using recurrent neural networks to process data with time-series features can often get better results. The LSTM network was constructed as follows:

Data processing: We extracted the data at fixed interval T to predict the data direction periodically. Thereafter, we obtained a matrix of $1 * T$ as the model input;

Determine the number of layers of neural network: We used a two-layer model with an LSTM layer and a fully connected layer to solve the nonlinearity and temporality problems in the prediction process;

Loss function: Here, we chose the Mean square error as the loss function.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(observed_i - predicted_i)^2, \tag{11}$$

Where $observed_i$ is the original data and $predicted_i$ is the calculated predicted data.

3.2. Solving the optimization problem based on GA

GA is a heuristic search method (Katoch, Chauhan, and Kumar 2021), which can obtain better optimization results faster than some conventional optimization algorithms when solving more complex combinatorial optimization problems(Ullah et al. 2021).We improved GA to make it more accurate for solving the goal programming problem while ensuring solvability.

Binary encoding: We first selecting the top N suppliers through quantitative analysis as the base selection. Then, coding these N suppliers by creating a $1 * N$ matrix to represent the N suppliers. In this matrix, each digit represents the selection of each supplier separately, with 0 representing not selected and 1 representing selected, such as $\underbrace{01001\cdots00101}_{1\times\mathrm{N}}$.

Constraint equation: To ensure that the GA can find the optimal solution. Therefore, restrictions were added to the population production to ensure that the produced individuals meet the basic requirements of the question. Moreover, adding constraints can significantly reduce the number of iterations of GA. Without the inclusion of constraints, the GA may be unsolvable. Constraint equation are defined as an order quantity to meet the production requirements.

$$M_{supplier} \times M_{quantity} \geq S, \tag{12}$$

$M_supplier$ is the coded supplier matrix and $M_quantity$ is the supply matrix.

Fitness function: The fitness function represents the degree of optimal degree of individuals in a population in GA, and it is used to distinguish between "good and bad" individuals. The fitness of an individual is generally determined by individual features. Therefore, it is defined as

$$f(x) = e^{-\sum_{i=0}^{N} M_i}, \tag{13}$$

Where M_i is the i-th term in the matrix $M_{supplier}$.

Binding Conditions: To ensure proper production, the production capacity provided by the various raw materials must meet the production capacity required by the company Cap. Therefore, additional constraints should be added to the GA. Assuming that loss (%) of the raw material is lost in transit, the supply capacity mathematical equation is defined as:

$$loss\% \times \left(\frac{sel \cdot mat}{l}\right) \geq \begin{pmatrix} Cap & & \\ & \ddots & \\ & & Cap \end{pmatrix}, \tag{14}$$

Where sel is the supplier's selection matrix, mat is the supplier's availability, and l is the original production conversion rate.

Select operation: The selection operation chosen for this scheme is the roulette selection method, and it is defined as

$$p_i = \frac{f_i}{\sum_{i=1}^{N} f_i}, f_i = \frac{k}{F_i} N, \tag{15}$$

Mutation: Random selection of chromosome X and its y-th gene α_{xy} is mutated according to a certain mutation probability to produce a new individual.It is defined as

$$a_{xy} = \begin{cases} [(1+e^{y-\frac{N}{2}})^{-1} + 1], r < 0.5 \\ [(1+e^{y-\frac{N}{2}})^{-1} + 0], r > 0.5 \end{cases}, \tag{16}$$

Where N is the length of the chromosome. $y = [x]$ is a Gaussian function and r is a random number between [0,1].

4 Experiments

In this study, we selected a supply chain of an enterprise. The production enterprise mainly produces construction and decorative panels, the raw materials used include wood fiber and other plant vegetal fiber materials, which can be divided into three types of A, B and C. The capacity of the enterprise is 28,200 cubic meters per week, and each cubic meter of product requires 0.6 cubic meters of raw material of category A, 0.66 cubic meters of raw material of category B, or 0.72 cubic meters of raw material of category C. The data provides the order quantity and the actual supply quantity of 402 suppliers.

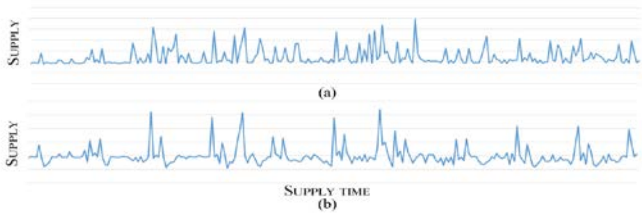


Fig. 2. Raw data (a), Preprocessed data (b).

4.1. Data preprocessing

First we processed the data for outliers and the result is presented in Figure 2. Comparing Figure 2 (a) and (b) indicates the periodicity of the preprocessed data is stronger.

4.2. Quantitative evaluation of suppliers

After feature processing, we used entropy-TOPSIS evaluation model to quantitatively evaluate 402 suppliers, and the results obtained are summarized in Table 1.

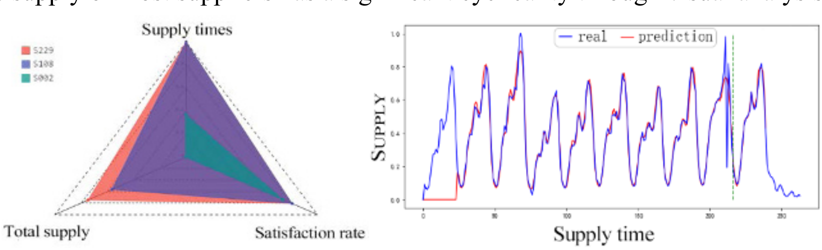
According to the composite score index derived from the entropy-TOPSIS method, S229 and others are the suppliers that can be stable to work with in the future years. Therefore, the top-ranked suppliers are the most important suppliers for this company. The feature indicators we extracted are displayed using radar plots (the larger the area enclosed, the more prominent the indicators are), as shown in Figure 3 (a), which also shows that suppliers with relatively high overall scores, such as S229 and S108, have more balanced indicators.

Table 1. Supplier quantitative evaluation results.

Ranking	1	2	3	4	5	6	7
Supplier number	S229	S108	S308	S151	S361	S140	S330
Comprehensive score index	0.641	0.61	0.6	0.571	0.57	0.566	0.562

4.3. LSTM prediction and GA to formulate the optimal scheme

In this study, we proposed a solution for the optimal ordering scheme without considering the premise of a sudden high order from the enterprise. For the raw data, we established that the supply of most suppliers has a significant cyclicality through visual analysis.



(a) (b)

Fig. 3. Comparison of features indexes of some suppliers (a), Supplier data fitting effect (b).

First, we used the LSTM model to predict the data. From Figure 3 (b), the prediction results are consistent with the actual and most of them have regularity. This ensures the feasibility of the subsequent purchase plan. Second, we used GA to develop the most economical raw material ordering scheme. The results are summarized in Table 2. Each column of the table represents the purchase plan in week n. For instance, $2634\ m^3$ at supplier $S140$ in week 1 and $1410\ m^3$ at supplier $S338$.

The results indicate that when choosing a supplier, preference is given to those with whom we have had a long and stable relationship in the past. The selection of suppliers and supply quantities is also shown in Figure 4, where the horizontal coordinates represent the supplier numbers and the vertical coordinates represent the advised supply quantities. Figure 4, we can see that suppliers such as S140, S151 and S229 have a high order quantity. Comparing their quantitative evaluation results indicates that these suppliers have relatively high overall score indices. Therefore, the quantitative evaluation results of suppliers have certain reference value.

Table 2. Best procurement scheme (part)

Supplier number	S140	S338	S229	S151	S086	S308	S139
Week 01	2634	1410	1680	4025	1068	688	6226
Week 02	4255	2010	1876	1720	1500	1162	1128
Week 03	1104	-	1139	2466	-	1002	358

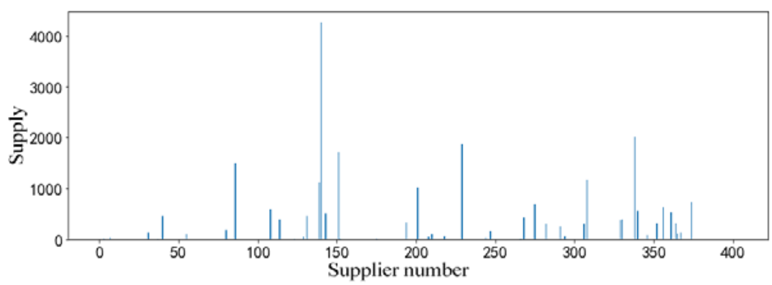


Fig. 4. Selection of suppliers and supply quantity.

5 Conclusion

In this study, we used entropy-TOPSIS, GA and LSTM neural network to solve the decision planning problems in enterprise raw material supply chain. First, we establish a supplier evaluation framework based on entropy-TOPSIS to evaluate the problem quantitatively. Second, we use the improved GA and LSTM models to find the suppliers that meet the production needs of the enterprise and to develop the optimal ordering scheme. Our model starts from the group searching, which has potential parallelism, and can compare multiple individuals at the same time. Thus, the problem can be solved more efficiently. The results indicate that our method can efficiently make a certain reference value for the future ordering plan of the company without considering the premise of the company's explosive order to some suppliers, and it can effectively quantify the supplier analysis; thus, the company can make a better choice.

This work was supported in part by the National Natural Science Foundation of China under Grant 61903090.

References

[1] H. Min, “Artificial intelligence in supply chain management: theory and applications,” *International Journal of Logistics: Research and Applications*, vol. 13, no. 1, pp. 13–39, 2010.

[2] S. Raman, N. Patwa, I. Niranjan, U. Ranjan, K. Moorthy, and A. Mehta, “Impact of big data on supply chain management,” *International Journal of Logistics Research and Applications*, vol. 21, no. 6, pp. 579–596, 2018.

[3] G. Baryannis, S. Validi, S. Dani, and G. Antoniou, “Supply chain risk management and artificial intelligence: state of the art and future research directions,” *International Journal of Production Research*, vol. 57, no. 7, pp. 2179–2202, 2019.

[4] X. Xu, X. Chen, F. Jia, S. Brown, Y. Gong, and Y. Xu, “Supply chain finance: A systematic literature review and bibliometric analysis,” *International Journal of Production Economics*, vol. 204, pp. 160–173, 2018.

[5] M. Ben-Daya, E. Hassini, and Z. Bahroun, “Internet of things and supply chain management: a literature review,” *International Journal of Production Research*, vol. 57, no. 15-16, pp. 4719–4742, 2019.

[6] R. Dash, M. McMurtrey, C. Rebman, and U. K. Kar, “Application of artificial intelligence in automation of supply chain management,” *Journal of Strategic Innovation and Sustainability*, vol. 14, no. 3, pp. 43–53, 2019.

[7] Y.-J. Lai, T.-Y. Liu, and C.-L. Hwang, “Topsis for modm,” *European journal of operational research*, vol. 76, no. 3, pp. 486–500, 1994.

[8] J. A. V´asquez, J. W. Escobar, and D. F. Manotas, “Ahp–topsis methodology for stock portfolio investments,” *Risks*, vol. 10, no. 1, p. 4, 2021.

[9] X. Huang, “Evaluation of vehicle handling stability based on interval topsis and entropy weights,” in *Journal of Physics: Conference Series*, vol. 2095, no. 1. IOP Publishing, 2021, p. 012055.

[10] M.-L. Shen, C.-F. Lee, H.-H. Liu, P.-Y. Chang, and C.-H. Yang, “Effective multinational trade forecasting using lstm recurrent neural network,” *Expert Systems with Applications*, vol. 182, p. 115199, 2021.

[11] H. Chen, S. Liu, R. M. Magomedov, and A. A. Davidyants, “Optimization of inflow performance relationship curves for an oil reservoir by genetic algorithm coupled with artificial neural-intelligence networks,” *Energy Reports*, vol. 7, pp. 3116–3124, 2021.

[12] S. Katoch, S. S. Chauhan, and V. Kumar, “A review on genetic algorithm: past, present, and future,” *Multimedia Tools and Applications*, vol. 80, no. 5, pp. 8091–8126, 2021.

[13] Z. Ullah, S. R. Naqvi, W. Farooq, H. Yang, S. Wang, D.-V. N. Vo et al., “A comparative study of machine learning methods for bio-oil yield prediction–a genetic algorithm-based features selection,” *Bioresource Technology*, vol. 335, p. 125292, 2021.