Steel logistics park layout based on market basket analysis and accessibility

Qianqian Yang[1], and *Lu* Qin[1,*]

[1]Beijing Jiaotong University, China

Abstract. In this study, we used market basket analysis and land accessibility for the layout of MG Steel Logistics Park. Due to the unreasonable cargo space layout, the outbound trucks in the MG Steel Logistics Park detour and transfer, which caused congestion. Using market basket analysis, this research took outbound trucks as the research object to measure the degree of correlation between cargo locations. Then, we measured the convenience of cargo locations by the accessibility and matched it with the steel outbound frequency. We established a model and got the layout allocation plan with the maximum degree of relevance and the minimum difference in cargo-location matching as the target. In the solution algorithm, we use roulette method and the elite retention strategy to improve the genetic algorithm. Finally, we compared the results before and after optimization. We found that the number of truck transfers was reduced by 33%, and the truck staying time was reduced by 21%.

1 Introduction

MG Steel Logistics Park is the largest one in Anhui Province, China, whose main task is to store and distribute steel. The inbound is determined by the production plan of the steel mill. This study only researches the outbound process, involving 104 cargo spaces and 60 types of steel. In this study, customers place orders online and arrange trucks offline to pick up steel. After analysis, an order often contains 1-10 SKUs, corresponding to several truck. By the way, the park arranges the outbound plan of steel materials for pickup trucks. An outbound order often contains 3-4 SKUs and positions.

Steel has some strict storage and loading requirements: only one type of SKU can be stored in one cargo space; when loading, the priority should be sorted according to the shape and characteristics of the steel, long>short>thick>thin.

Due to the unreasonable layout of cargo spaces, truck transfers from one yard to another for too many times, which results in serious congestion in the park. The excessive truck staying time, affects the work efficiency seriously. So, this study places highly related steels nearby to reduce truck transfer and detour times. At the same time, the steel materials that are frequently out of the warehouse are arranged in a cargo location with higher accessibility to the exit to reduce the staying time of the trucks in the park.

---

* Corresponding author: lqin@bjtu.edu.cn

2 Methods

2.1 Relevance

The layout of the logistics park often adopts the system layout planning (SLP) and its improvement methods proposed by Richard Muther. These methods usually use the total volume of logistics as the basis to measure the degree of relevance. However, it is the total number of outbound trucks and the degree of interconnection that determine the congestion level of the park. Therefore, the total volume of logistics is not representative and it is necessary to research the degree of steel correlation reflected by outbound trucks.

The most appropriate method is market basket analysis. As a method to find the possibility of jointly purchasing products, market basket analysis widely applies association rules [1]. Association rules have been successfully used in product classification decision [2], profit mining [3][4], credit card business [5], market analysis [6] and other fields. They are also used in the field of layout [7] [8].

MBA has three parameters which are support, confidence, and lift. Support is a proportion of event B because of event A. Confidence is a probability event B happened because of event A dependently. Lift is a probability of event B happened because of event A independently [9].

In this study, each truck corresponds to "a customer". The warehouse receipt of each truck corresponds to a "market receipt". The different steel products in the warehouse record correspond to "retail goods". This study analyzes the outbound documents of all trucks by Apriori association rules, and measures the degree of steel association by confidence.

2.1 Accessibility

Accessibility refers to the degree of difficulty from one location to another, and it can be used in land use planning [10]. Steel cargo space accessibility refers to the degree of difficulty of the whole process of trucks entering and leaving the park, including unloading/collecting goods at the designated cargo space. There are topological methods and geometric network methods to measure accessibility.

In this study, the spatial syntax in the topology method is adopted. The integration degree of each cargo location is obtained by the field of view analysis. The integration degree reflects the agglomeration or dispersion between a space and all other spaces in the system. The larger the value, the stronger the accessibility. When it is between 0.4 and 0.6, the layout is scattered. The calculation formula for integration degree is:

$$I = RA_m = \frac{2(MD_m - 1)}{n - 2} \tag{1}$$

Among them, MD_m is the average depth value, which represents the shortest distance from a certain node to all other nodes. The calculation formula is as follows:

$$MD_m = \frac{\sum_{k=1}^{n} dmk}{n - 1} \tag{2}$$

In the formula, d_{mk} is the shortest distance between any two points m and k in the connection graph, and n is the number of all nodes in the connection graph.

2.3 The degree of steel outbound

The degree of steel outbound is directly proportional to the outbound frequency and quantity. First, the deviation standardization method is used in the dimensionless normalization of the

inbound and outbound frequency and the total volume of goods entering and leaving the warehouse. The formula is as follows:

$$X^* = \frac{X - min}{max - min} \tag{3}$$

The formula for calculating the degree of steel leaving the warehouse is :

$$S_i = \alpha \cdot Fre_{out} + \beta \cdot Num_{out} \tag{4}$$

Notation:

Si represents the degree of steel outbound; α, β respectively represent the weighting coefficients of the frequency and quantity of outbound, $\alpha+\beta=1$, where α is 0.7; Fre_{out} represents the frequency of outbound; Num_{out} represents the quantity of outbound.

3 Model and mathematics

In this study, the research objective consists of two objective functions. *F1* represents the total correlation between the steel products after the location layout is completed, and *F2* represents the balance between the degree of steel out of the warehouse and the convenience of the location.

$$maxF1 = \sum_{i,j=1, i \neq j}^{S} \sum_{p=1}^{P} \frac{R_{ij}}{d_{ij}} \cdot X_{ip} \tag{5}$$

$$minF2 = \sum_{i=1}^{S} \sum_{p=1}^{P} \left| \frac{E_i}{E_p} - 1 \right| \cdot X_{ip} \tag{6}$$

$$maxF = \frac{F1}{F2} = \frac{\sum_{i,j=1, i \neq j}^{S} \sum_{p=1}^{P} \frac{R_{ij}}{d_{ij}} \cdot X_{ip}}{\sum_{i=1}^{S} \sum_{p=1}^{P} \left| \frac{E_i}{E_p} - 1 \right| \cdot X_{ip}} \tag{7}$$

Restrictions:

$\sum_{i=1}^{S} X_{ip} = 1$,it means that only one kind of steel is stacked in each position.

$\sum_{p=1}^{P} X_{ip} = N_i$,it indicates that the layout result must meet the number of cargo spaces required for each steel.

Meaning of symbols:

S represents the set of steel, i and j respectively represent any two SKUs; *P* represents the set of cargo locations, and *p* represents any of them; *Rij* represents the degree of correlation between steel *i* and *j*; Dij represents the distance between steel *i* and *j* corresponding to the cargo space under a certain layout plan; *Xip* indicates whether the steel *i* is placed on the cargo position *p*, if it is 1, if not, it is 0; *Ei* represents the score of the frequency of steel in and out of warehouse; *Ep* represents the score of the accessibility of the cargo location; *Ni* represents the quantity of cargo space required for steel *i*.

4 Improved genetic algorithm

4.1 Basic parameter setting

This model is solved by an improved genetic algorithm. The population size is set to 50, the evolution termination iteration algebra is set to 500, the crossover probability is set to 0.5, the mutation probability is set to 0.05, and the fitness function value is the objective function value *F* of the model.

4.2 Chromosome design

Since the number of steel SKUs is not equal to the number of locations, it is important to design chromosomes. Number each cargo location from each yard, and number the steel SKU.

Chromosomes are represented by a 0-1 matrix with rows less than columns. The row number is the number of the steel SKU, and the column number is the location number. The sum of the row elements is equal to the number of locations required for the steel, and the sum of all the column elements is equal to 1. Crossover, inheritance, and mutation are operations on column numbers——column number is gene.

4.2 solution process

The solution process design is as follows:

(1) Generate initial population

Randomly generate 50 permutations from 1 to 104 as the initial population.

(2) Crossover& Inheritance

In the population, another chromosome is randomly selected for each chromosome and the two are crossed. The corresponding positions of the two parental chromosomes have a probability of crossover with a probability of 0.5. For genes that are repeated after the crossover, fill in the missing genes with the missing genes. A total of 100 individuals were obtained from the parents and offspring through inheritance.

(3) Elite retention strategy

Calculate the fitness values of a total of 100 individuals of the parent and offspring, and sort them in descending order, and keep the top 25 individuals with the highest fitness.

(4) Roulette method

Using the total fitness of 25 individuals as the roulette wheel, and the proportion of each individual's fitness to the total fitness as the probability, 50 chromosomes are reselected.

(5) Mutations

With a probability of 0.05, select the mutated individuals among 50 individuals, and randomly select two genes for these individuals to exchange, and a new generation of population will be produced.

(6) Determine whether the convergence condition is met, if it is met, end the loop, if it is not met, return to (2).

4.3 Rresult

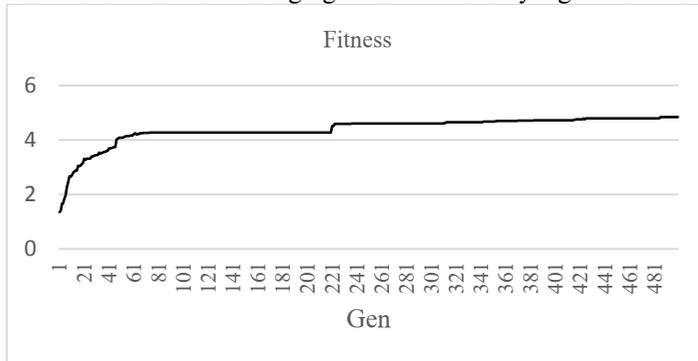The curve of fitness function value changing with evolutionary algebra is as follows:

Fig. 1. Fitness function value.

The layout result from chromosome transcoding is shown in the figure below.

**Table 1.** Excerpts from layout results.

Steel Number	Cargo Number
1	67
2	4/9/12/16/19/20/26/28/29/38
3	1/46
4	68
5	34/94/99/104
6	2/11
7	69/78
8	43
9	53
10	21/25/27/30/100
11	10/76
12	72
13	41
14	54
15	13/22/23
16	6/57
…	…

5 Optimization effect analysis

We take the data of the busiest day of the year and use two indicators, the number of truck transfers and the truck stay time, to simulate and compare the optimized layout and the initial layout.

(1) Initial layout

Before optimization, the total time of all trucks staying in the park was 3329 minutes, the number of outbound trucks was 288, and the average stay time of trucks was 11.6 minutes. The truck made a total of 68 transfers.
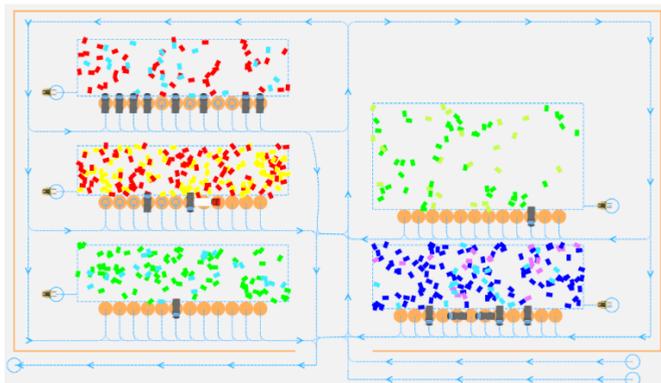
Fig. 2. Simulated image of original layout.

(2) Optimized layout

After optimization, the total time for all trucks to stay in the park is 2751 minutes, which is about 21% optimized. the number of outbound trucks is 278, and the average stay time of the trucks is about 9.9 minutes, which is about 17% optimized. The truck made a total of 51 transfers, a total of 17 reductions, an optimization of 33%.
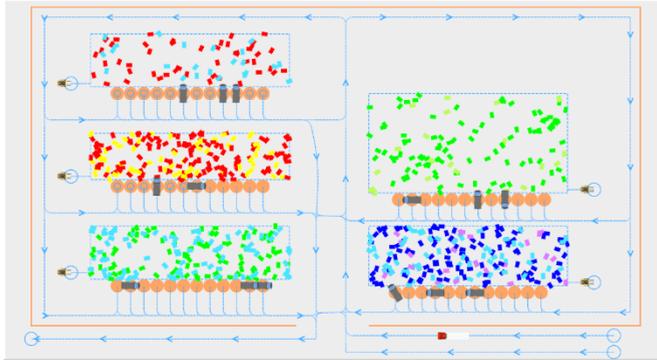
Fig. 3. Optimized simulation image of layout.

References

1. R. Agrawal, T. Imieliński, A. Swami,Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data. SIGMOD 93., ACM, New York, NY, USA (1993), pp. 207-216

2. T. Brijs, G. Swinnen, K. Vanhoof, G. Wets,Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. KDD '99, ACM, New York, NY, USA (1999), pp. 254-260, 10.1145/312129.312241

3. K. Wang, S. Zhou, J. Han,Springer Berlin Heidelberg, Berlin, Heidelberg (2002), pp. 70-87

4. R.C.W. Wong, A.W.C. Fu, K. Wang,Third IEEE International Conference on Data Mining (2003), pp. 371-378

5. R.C. Wu, R.S. Chen, C.C. Chang, J.Y. Chen,IEEE Comput. Soc. (2005), pp. 39-40

6. G.S. Linoff, M.J.A. BerryData Mining Techniques: For Marketing, Sales, and Customer Relationship Management(3rd), Wiley (2011)

7. Siana Halim, Tanti Octavia, Christian Alianto,Procedia Computer Science,Volume 161,2019,Pages 623-629,ISSN 1877-0509

8. Elif Ozgormus, Alice E. Smith,Volume 139,2020,105562,ISSN 0360-8352

9. Ledolter, J. (2013) Data Mining and Business Analytics with R, 1st Edition, John Wiley & Sons,Inc., New Jersey

10. Jonathan Levine,Transportation Research Part D: Transport and Environment,Volume 83,2020,102309,ISSN 1361-9209