



Research on optimization of e-commerce warehouse storage location assignment based on improved NSGA-II algorithm

Xinqi Wu^{1,*}, Juping Shao¹ and Yanan Sun²

¹ School of Business, Suzhou University of Science and Technology, Suzhou, Jiangsu, 215009, China

² Anwood Logistics System Co. Ltd, Suzhou, Jiangsu, 215021, China

SUMMARY: *The rising global e-commerce transactions and the sharp increase in the variety and quantity of goods in e-commerce warehouses inevitably bring new challenges to the storage location assignment. This paper takes movable shelves as the research object, and proposes a method of space allocation based on commodity correlation, while considering the quantity - distance correlation of goods. Aiming at optimizing the correlation of goods and the stability of shelves, an integer programming model was constructed, and an improved NSGA-II algorithm was designed to solve the problem by introducing a learning mechanism combined with the characteristics of the model. Finally, through concrete examples, it is proved that the storage location assignment model and algorithm proposed in this paper can effectively optimize the value of the objective function, improve the picking efficiency by reducing the average handling frequency of the shelves, and provide corresponding management suggestions for the effective operation of the movable shelf storage system.*

KEYWORDS: *E-commerce warehouse; Movable shelf system; Storage location assignment problem; Improved NSGA-II algorithm*

1 Introduction

With the continuous updating of Internet technology and the continuous evolution of people's consumption concept, e-commerce is experiencing rapid growth in the world. According to the Global E-commerce Report 2023^[1] released by the United Nations Conference on Trade and Development, as of December 2023, about 2.64 billion people in the world have completed at least one online purchase, and the total transaction value of the global e-commerce market has reached 5.82 trillion US dollars, which continues to rise compared with previous years. The continuous rise in the global transaction volume of e-commerce inevitably leads to a sharp increase in the order quantity, type and quantity of goods in e-commerce warehouses. The original storage location assignment scheme may soon become ineffective, and warehouses need to constantly adjust the space planning to adapt to the rapidly changing inventory structure. Such frequent adjustment increases the complexity of storage location assignment. The warehouse management system is required to have stronger dynamic response ability, which brings higher difficulties and challenges to the operation and development of the warehouse. In this context, whether e-commerce enterprises can withstand the test of orders with the characteristics of "small batch, multi-batch, timeliness", and adopt effective storage location assignment schemes to improve the efficiency of order picking is a problem of practical

*wuxinqi0715@163.com

<https://doi.org/10.65102/is20261146>

significance.

This paper is divided into six chapters. The first chapter is Introduction, including the research background and practical significance. The second chapter is Literature Review, summarizing the existing studies and pointing out the differences between this study and the existing studies. The third chapter is Problem Description and Modeling, describing the problem studied in this paper, determining the optimization objectives, and finally establishing a multi-objective optimization model under the assumptions and constraints. The fourth chapter is Algorithm Design, designing an improved NSGA-II algorithm according to the characteristics of the model to solve the model better. The fifth chapter is Instance Validation, combining the model with the actual e-commerce warehouse data, and using the improved NSGA-II algorithm to solve the best storage location assignment scheme. The sixth chapter is Conclusions, summarizing the research content of this paper and looking forward to the future research.

2 Literature Review

Specifically, storage location assignment is a process for enterprises to rationally allocate the storage location of goods according to the characteristics of goods, the needs of users and other related factors[2]. Effective storage location assignment is a key decision-making problem in warehouse operation[3]. By optimizing storage location assignment, storage space can be rationally allocated according to factors such as the size, weight and turnover rate of goods, so as to avoid excessive congestion or waste of warehouse space and improve the overall storage capacity of the warehouse; according to the specific needs of the order and the characteristics of the goods, the goods with high shipment frequency can be placed near the export position, so as to enhance the screening speed of the order and improve the rapid response to the order; it can reduce the energy use of the warehouse (such as electricity and transport fuel for equipment such as forklifts), while also reducing carbon emissions, which is very beneficial for enterprises to achieve sustainable development goals.

Traditionally, the problem of storage location assignment is mainly based on the picking mode of "person-to-goods", and now the research in this area has been quite rich[4]. In the "person-to-goods" picking mode, the picker will go to the corresponding shelf or storage area according to the order demand and the specified cargo location information, find the required goods and complete the selection. The model built is usually to optimize the travel distance or travel time of the picker[5-8]. For example, Koch[6] introduced a grouping genetic algorithm to group different customer orders into pick orders in order to minimize the total length of all warehouse trips. In order to further reduce the demand response time, Jiang[7] designed a novel picking strategy, which divided orders and assigned them to different pickers according to vehicle routes, collected all the order fragments, and delivered them to customers.

However, in the "person-to-goods" picking mode, the shelf is fixed. Therefore, when dealing with small batch, multi-batch, time-sensitive orders in the e-commerce environment, it will significantly increase the movement time of pickers, increase labor costs, reduce the efficiency of picking, and hinder the operation and development of enterprises[9-11]. Based on this, the "goods to person" picking mode came into being. Figure 1 shows the specific workflow of the picking mode. The "goods to person" picking model relies on automated tools (such as AGVs) to move the shelves to the picker, and when the picker finishes the picking job, the AGV will transport the shelves back to their original location for storage.

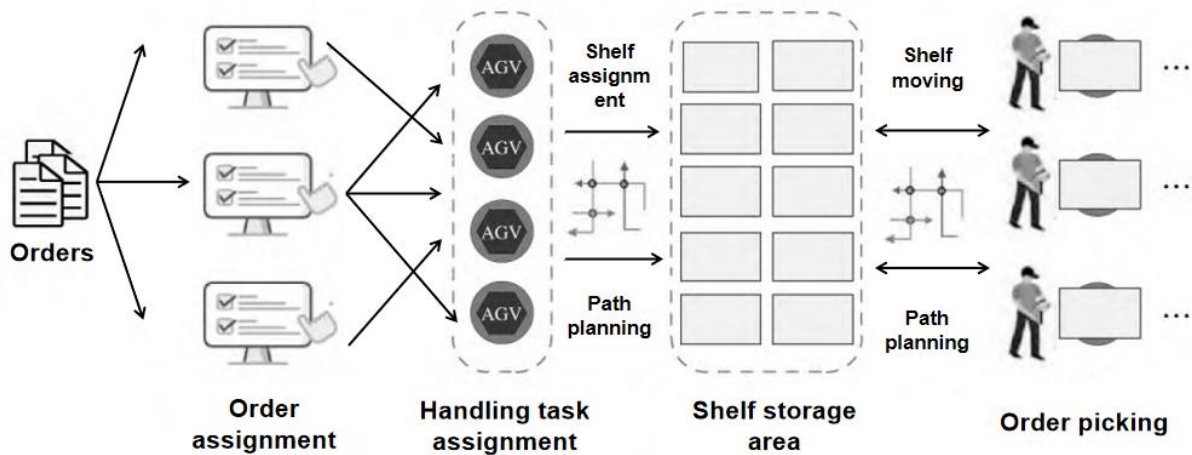


Figure 1: Specific workflow of the "goods to person" picking mode

Since each robot can only move one shelf at a time, given a batch of orders to be picked up and a limited number of robots, the number of times the shelf is handled greatly affects the efficiency of order picking, and the commodity storage allocation scheme determines the number of shelf handling. Figure 2 shows two orders to be selected and three storage allocation schemes. Plan 1 stores A and B products in different shelves; in Plan 2, both shelves store 2 A's and 2 B's; in Plan 3, one shelf stores 3 A's and 1 B's, and the other shelf stores 1 A's and 3 B's. Three storage allocation schemes to complete these two order picks require moving the shelves 4, 4 and 2 times respectively. Plan 3 effectively improves the picking efficiency through rational storage allocation. In actual operation, in the face of large-scale product mix and quantity relationship reflected by massive orders, storage allocation needs to determine the type and quantity of goods stored on each shelf to meet the diversified order picking needs as much as possible, thus minimizing the number of shelf movements required for order picking.

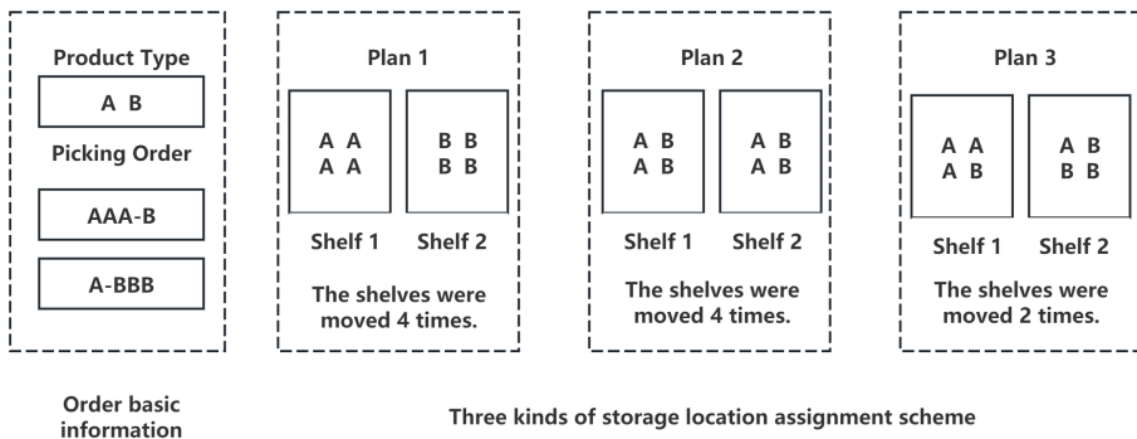


Figure 2: Examples of commodity storage location assignment.

Existing studies about "goods to person" picking mode mainly focus on two levels of model construction and algorithm design:(1) In terms of model construction, scholars often model this problem as an integer programming model, and choose to enhance the relevance of products on the shelves[12-16], optimized frequency of incoming and outgoing storage[17-20],

and minimize the picking time[21] or picking cost[22] as the optimization goal. Besides, Xiao[12] established an optimization model for minimizing the order picking path and maximizing the picking rate based on BOM information. Xiang Shi[29] studied a unique shelf scheduling problem with multiple types and multiple storage locations, determined the retrieval order of the shelves, and assigned a storage location to each shelf after visiting the pick station. Kim et al.[30] studied the possibility of placing a commodity on multiple shelves and proposed an optimization heuristic method based on commodity correlation degree to solve the problem of space allocation.(2)In terms of algorithm design, considering the NP-hard characteristics[23] of the problem and the large-scale nature of real life, there are limitations to the use of precise algorithms, and heuristic algorithm has become a common solution technology. The algorithms commonly used to solve the problem of storage location assignment include genetic algorithm, simulated annealing algorithm, ant colony algorithm, particle swarm algorithm and tabu search algorithm. Lin et al. [36] used genetic algorithm to calculate cargo location allocation and established a multi-objective optimization model considering the determination of retrieval time and frequency. The results show that the search ability of GA is greatly improved under the condition of frequent entry and departure. Muppani et al. [26] used the simulated annealing algorithm to solve the linear optimization model of space allocation controlling space utilization and picking cost, and proved that the simulated annealing algorithm was more accurate than the dynamic programming algorithm through practical data cases. Chen et al. [37] proposed a kind of automatic retrieval and storage space allocation problem with time window, and proposed an improved tabu search algorithm to solve this problem, and the results proved that the solution obtained by this algorithm was of higher quality. In addition, representative algorithms also include Xiang et al.'s variable neighborhood search algorithm[24], Guan Mengchen and Li Zhenping' s improved genetic algorithm[25], Mirzaei et al. 's iterative greed algorithm[21], Li Zhenping et al.' s single-parent legacy algorithm[22], as well as Yang P's Tabu search algorithm[27] and so on. At other levels, Hu[28] proposed a three-stage space allocation method of "association network construction→association network analysis→association network clustering" based on the idea of reducing the problem solution space, and mapped the clustered commodity network into the corresponding space placement of the shelves to quickly and efficiently generate the space allocation scheme. Aiming at minimizing the selection cost, Ding[31] proposed a three-stage iterative search and selection method based on association network for efficient solution to the challenges faced by the double selection system, such as complex commodity correlation relationship and huge solution space.Wang[32] established shelf heat and correlation generation models based on customer orders, and based on these two attributes, established a bi-objective integer programming model considering the minimization of robot heavy load and no-load traveling distance, and designed a heuristic solution algorithm for tabu search. Zhai[38] studied the commodity storage allocation problem based on the moveable shelf selection model, constructed an integer programming model aiming at minimizing the total number of shelf movements, and designed a variable neighbor-tabu search algorithm conforming to the characteristics of the research problem.

The differences between this paper and the existing researches are mainly reflected in two aspects: research content and solving algorithm. In terms of research content, existing scholars in the study of "goods to person" picking mode under the storage location assignment problem, usually directly assumed in a particular warehouse background, the initial state of the shelves is empty, and a commodity can only be stored in a particular shelf, that is, the static, one commodity one space of the storage location assignment problem. However, it does not meet the actual operation and cannot meet the sustainable development of the enterprise. In actual

operation, goods are often scattered on multiple shelves, and the type and quantity of goods in the warehouse are not fixed, and it is necessary to regularly replenish, remove some unsalable goods and put new goods on the shelves, which is a dynamic process. And in terms of solving algorithm, this paper selects NSGA-II algorithm to solve the model, and introduces learning mechanism combined with the characteristics of the cargo location allocation model to improve and innovate the algorithm. In summary, this study mainly focuses on the dynamic and one commodity multiple spaces of the storage location assignment problem in the e-commerce environment. As shown in Figure 3, a multi-objective integer programming model is constructed considering multiple factors such as commodity quantity -- distance correlation degree and shelf stability, and the improved NSGA-II algorithm is adopted to solve the problem. Meanwhile, the optimization performance of the algorithm is analyzed. Finally, the corresponding management enlightenment is obtained.

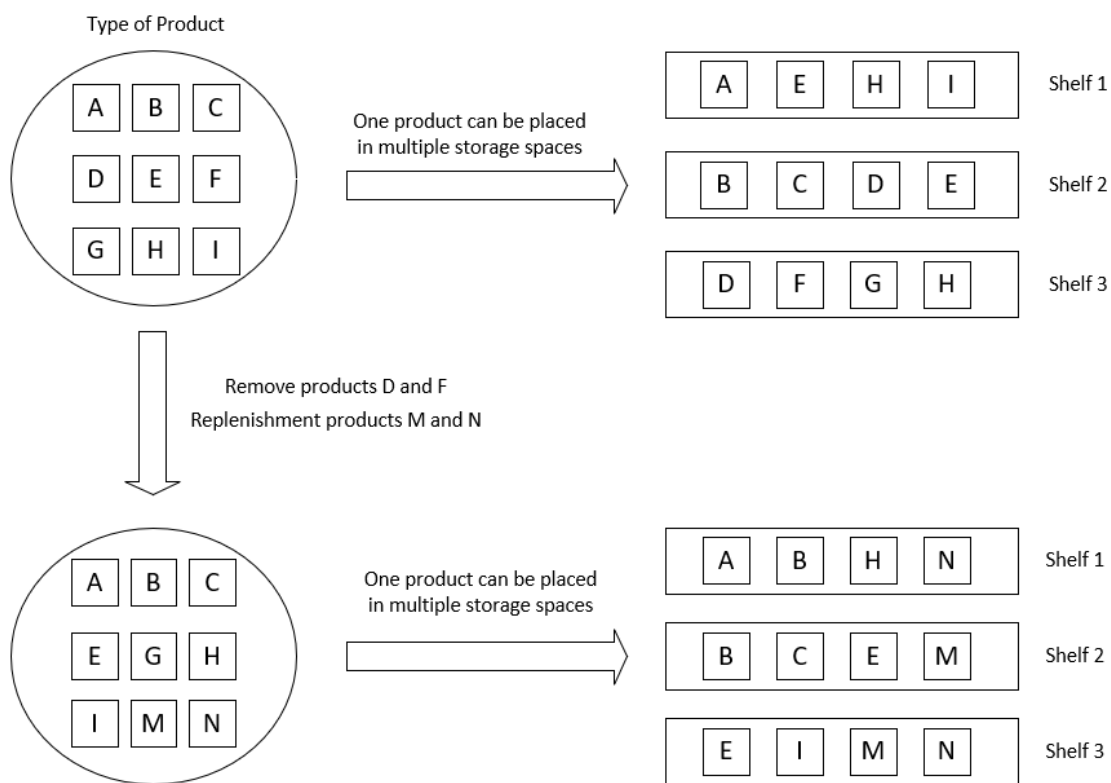


Figure 3: Dynamic storage location assignment problem.

3 Problem Description and Modeling

3.1 Description of the issue

Optimization of the assignment of storage locations in e-commerce warehouses as an important link to reduce the time needed to operate, minimize the moving distance in and out the warehouse, and utilize the entire amount of storage space[33], the specific problem can be described as follows: an e-commerce warehouse has a total of P kinds of commodities, which are stored in S shelves, and each shelf has N floors, with Q storage spaces in each floor. Each storage space of a shelf allows only one type of commodity to be stored, but the same commodities can be stored in different storage spaces. It is understood that an e-commerce

company will receive T orders for a specific region in the future period of time, and the types of goods in these orders can be found in the warehouse. Therefore, how to reasonably allocate storage locations for these goods, that is, to clarify the type and quantity of goods on the shelf, in order to reduce the cost of picking, improve the efficiency of picking, and ensure the stability of the shelf, is the core issue of this study.

Due to the complex correlation relationship between goods and shelves, this paper will introduce the concepts of commodity correlation degree, spatial dispersion degree and shelf stability. Among them, the commodity correlation degree includes the double-layer relationship between quantity and distance, the quantity relationship refers to the ratio between the quantity of orders containing two items and the total quantity of orders in the same batch of orders, and the distance relationship is the spatial dispersion degree. A product in the warehouse will be stored in multiple cargo locations. In this paper, the spatial dispersion of a product is represented by the sum of the spatial linear distance between all the coordinates of the product and the mean coordinates of the product, so as to minimize the spatial distance between the same product and the goods with strong quantitative correlation. The stability of the shelf means that when the storage space is allocated, the strategy of “light on top and heavy on bottom” must be followed to ensure that the center of gravity of the shelf is always maintained at a low level. This mathematical model can be abstracted as the sum of the product of the mass of the goods in each storage space and the floor on which it resides. This paper takes the optimization of commodity correlation degree and shelf stability in the warehouse as the research objective, mainly for the following three reasons: (1) Reduce shelf movement and improve warehouse security. If the less related commodities are kept in the same place, and the shelf stability of the goods is low, which largely means that the goods contained in an order are not in the same place, then the picker needs to complete the selection of an order, the number of times that the shelf needs to move compared to the number of times that the related goods are stored together will be increased. The loss of shelves will also increase accordingly. At the same time, the center of gravity of the higher shelves in the handling process, the risk of the warehouse will inevitably be increased. (2) Reduce unnecessary waste of time. If the goods of the same order are not stored in one place, the picker in the picking may appear more than one shelf at the same time to visit the picking table situation, resulting in warehouse congestion, queuing and waiting for a longer period of time, and ultimately unable to improve the picking efficiency. (3) Reduce picking costs. If the warehouse is allocated according to the degree of commodity correlation, the number of times the shelves are moved can be reduced, thus reducing the depreciation cost of the shelves. At the same time, the reduction of order picking time means that the picking efficiency, inventory turnover rate has been improved, also contributing to the reduction of picking costs.

In order to have a clearer understanding of the problem and model, the following hypotheses are proposed:

Assumption 1: The initial state of the warehouse shelves is not empty, so the research problem of this paper can also be described as the storage location assignment problem of replenished goods (replenished goods include original goods on the shelves and new goods). And the types and quantities of goods that need to be replenished over a period of time, as well as the types of goods that do not need to be replenished on each shelf and the empty storage on the shelf are known.

Assumption 2: In the order picking process, the number of orders to be picked and the commodity information contained in each order are known, and the types and quantities of commodities contained in each order can be split and picked.

Assumption 3: Each storage space is only allowed to store a commodity, but the same varieties of commodities can be stored in different storage spaces, i.e., assuming that there is no correspondence between the commodity and the storage space, and that any one commodity can be stored in any one of the spare storage spaces on any shelf.

Assumption 4: The capacity of the storage spaces on all shelves in the warehouse is the same, but the upper limit of the storage capacity of each compartment for each product is different due to the different size specifications of each product. Products are stored on shelves in full box form.

Assumption 5: The pickup truck runs at a uniform speed in the horizontal and vertical directions, without considering acceleration, and the traveling route is S-type, and the export of the goods is the coordinate origin.

3.2 Symbol Design

The meanings of the relevant parameters and variables contained in this article are shown in Table 1 and Table 2.

Table 1: Related parameters in the model of this article.

Parameters	Meaning
S	Total number of rows of shelves, that is, the total number of shelves, $t = 1, 2, \dots, S$
t	Shelf code, $t = 1, 2, \dots, S$
L	Number of shelf columns, that is, the storage number of each layer, $l = 1, 2, \dots, L$
N	Number of shelf levels, ground level is the first level, $n = 1, 2, \dots, N$
P	Total commodity category
i	Commodity category number, $i = 1, 2, \dots, P$
q_i	Number of storage spaces required to store the i th commodity
k	The k th storage space of the i th product, $k = 1, 2, \dots, q_i$
T	Total number of orders
T_i	Orders containing product i
r_{ij}	The degree of correlation between two goods i and j
$C_i(x,y,z)$	Mean coordinates of the i th product
$C_j(x,y,z)$	Mean coordinates of the j th product
M_{ik}	Weight of all products in the k th storage space of the i th product
d	Distance between product i and product j at picking time
d_1	Spatial dispersion between products of the same type
d_2	Spatial dispersion between different products

Table 2: Related decision variables in the model of this article.

Decision variables	Meaning
X_{ik}	Horizontal coordinate of the k th storage space of the i th product, i.e., the sorting number in which the storage space is located.
Y_{ik}	Vertical coordinate of the k th storage space of the i th product, i.e., the number of the column in which the storage space is located.
Z_{ik}	Height coordinates of the k th storage space of the i th product, i.e., the number of the layer in which the bay is located.

3.3 Modeling

3.3.1 Calculation of Commodity Relevance

It is known that r_{ij} denotes the degree of correlation between commodity i and commodity j , when $i \neq j$, $r_{ij} = \frac{T_{ij}}{T_i + T_j - T_{ij}}$; when $i = j$, $r_{ij} = 1$, where T_{ij} denotes the number of orders that contain both commodity i and commodity j , T_i is the number of orders that contain commodity i , and T_j is the number of orders that contain commodity j .

3.3.2 Calculation of Spatial Dispersion

1. Calculation of the degree of spatial dispersion between the same products.

The location in the warehouse is regarded as the point in the coordinate system, and the degree of dispersion of the product in space can be expressed by the sum of the spatial distance between the coordinate of each product and the mean coordinate. Among them, the mean coordinate of the i th product is:

$$C_i(x, y, z) = \frac{1}{q_i} \sum_{k=1}^{q_i} [x_{ik}, y_{ik}, z_{ik}], \quad (1)$$

Therefore, the degree of dispersion between all products can be expressed as:

$$d_1 = \sum_{i=1}^P \sum_{k=1}^{q_i} \sqrt{(x_{ik} - C_i(x))^2 + (y_{ik} - C_i(y))^2 + (z_{ik} - C_i(z))^2}, \quad (2)$$

2. Calculation of the degree of spatial dispersion between different products.

In this paper, the traveling distance is used instead of the linear distance in space to measure the degree of spatial dispersion, and the discrete distance between the i th product and the j th product is expressed as:

$$d = \left| (1 + n - 1) \times \left(\frac{C_i(x)}{2} - \frac{C_j(x)}{2} \right) \right| + |(C_i(x) - C_j(x))| + |(C_i(y) - C_j(y))| + |(C_i(z) - C_j(z))|, \quad (3)$$

The degree of dispersion of the i th product and the j th product is calculated by the product of travel distance and the degree of correlation, which is specifically expressed as:

$$d_2 = \sum_{i=1}^P \sum_{j=1}^P r_{ij} \times d, \quad (4)$$

Since the e-commerce warehouse has a strong correlation in order demand, the products with strong correlation need to be placed close to each other in the allocation of cargo space, so as to improve the picking efficiency, which needs to reduce the discretization between products. So the mathematical model of product relevance and spatial dispersion can be expressed as:

$$\text{MinD} = d_1 + d_2, \quad (5)$$

3.3.3 Calculation of Shelf Stability

The warehouse should allocate space according to the "light on top and heavy on bottom" principle to help ensure the stability of the shelf load, so that the center of gravity of the shelf is low, that is, the sum of products of the quality of the goods on each storage position and the layer where it is located is the smallest, and the stability of the shelf is the best. The specific mathematical model of shelf stability can be described as:

$$\text{MinF} = \sum_{i=1}^P \sum_{k=1}^{q_i} M_{ik} \times z_{ik}, \quad (6)$$

Based on the above variables and conditions, the following mixed integer programming model can be used to represent the storage location assignment problem discussed in this paper.

Objective function:

$$\text{Minf}_1 = D, \quad (7)$$

$$\text{Minf}_2 = F, \quad (8)$$

Constraints:

$$\begin{cases} 1 \leq x \leq t \\ 1 \leq y \leq l, \text{ and } x, y, z \text{ are positive integers} \\ 1 \leq z \leq n \end{cases} \quad (9)$$

The objective functions (7) and (8) represent the optimization of commodity relevance, spatial dispersion and shelf stability; the constraints (9) indicate that in the allocation of cargo space, it must be ensured that the values of x , y , and z must not be greater than the maximum of the number of rows, columns, and floors, and the coordinate values of the cargo space must be positive integers.

4 Algorithm Design

4.1 NSGA-II Algorithm Overview

NSGA-II algorithm is an improved version proposed by Deb et al in 2001 based on NSGA (No-dominated Sorting Genetic Algorithm)[34]. It is a fast non-dominated sorting multi-objective optimization algorithm with elite retention strategy. Detailed execution steps of NSGA-II algorithm are shown in Figure 4:

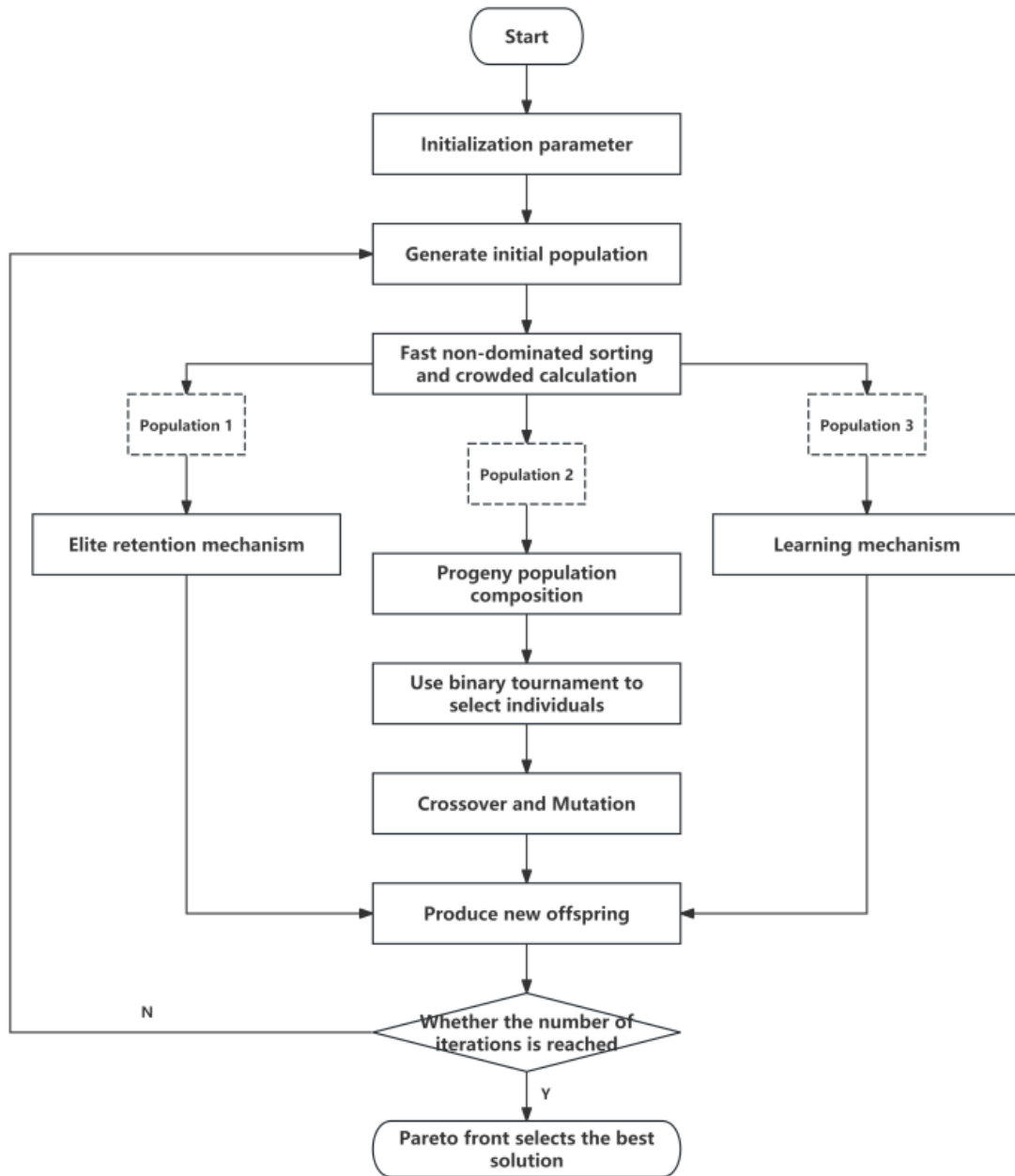


Figure 4: Detailed execution steps of NSGA-II algorithm.

A significant advantage of the NSGA-II algorithm is that it can rank individuals through rapid non-dominated sorting and generate Pareto frontiers. We take the best solutions of Pareto frontiers to construct the Pareto solution set, and use a hierarchical evolutionary strategy to determine the crowding distance between each Pareto frontier. Finally, solutions with better Pareto rank and larger crowding distance are selected as the child population, because this can better protect the diversity of the population and avoid the algorithm falling into the local optimal solution [35].

4.2 Improved NSGA-II Algorithm Solution

4.2.1 Coding of Storage Space Information and Population Initialization

In order to map the specific information of the storage space allocation scheme into genetic coding, this paper adopts integer coding, and the chromosome consists of storage space index number together with the cargo category number. If there are M goods, a random number is used to generate O individuals, each chromosome contains genes, each chromosome corresponds to a storage space allocation scheme, the specific form is shown in Figure 5. In the figure, goods 1 stored in the 3rd storage space in the 2nd floor of shelf 1.

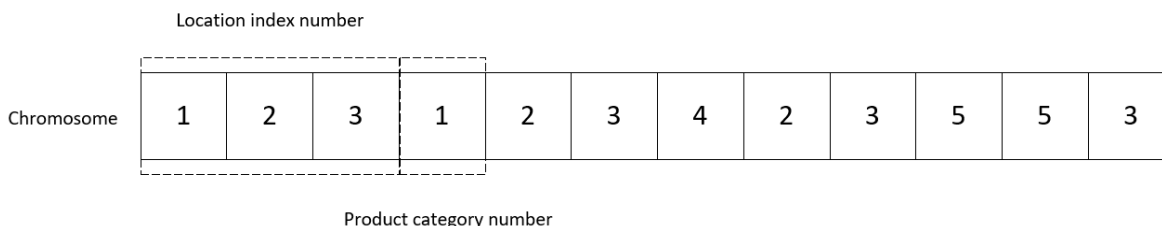


Figure 5: Chromosome coding method of storage location assignment.

Generating the initial population is to generate the initial chromosome population whose total number is Num under uncertainty. In this paper, the Rand function is used to randomly generate the above number of individuals to ensure that the generated chromosome meets the capacity requirements of the warehouse of this e-commerce enterprise.

4.2.2 Calculation of Fitness

The objective function in the model is typically used to establish the fitness function, in this paper, minimization serves as the objective function. The fitness function can be found as follows:

$$f_i = NN - Y_i \quad i = 1,2,3 \dots \text{Num}, \tag{10}$$

where Y_i is the objective function of the chromosome and NN is a large positive number, in which case the fitness will not be negative.

4.2.3 Fast Non-dominated Sorting and Congestion Distance Calculation

Suppose that the size of the population is of size I, the number of individuals in the population that are in the position of dominating individual i is n_i , and the set of individuals that are dominated by i is S_i . First find all the individuals in the population with $n_i = 0$ and save them into the current set F_1 . Then for each individual in the current set F_1 , search for each individual q in the set of individuals S_i that is dominated by i , and save the individual q to set F_2 if $n_q - 1 = 0$. Record the individual from the current set F_1 as the first non-dominant individual, set F_2 as the current set, and repeat the above steps until all individuals of the entire population species have been graded.

The congestion distance of an individual can be expressed as the sum of the differences in the objective function values of the two individuals adjacent to that individual in each optimization direction. The crowding distance of all boundary individuals is set to infinity, and for all intermediate individuals, the formula is as follows:

$$i_d = \sum_{j=1}^N |f_j^{i+1} - f_j^{i-1}|, \quad (11)$$

where i_d denotes the congestion distance of individual i , and f_j^{i+1} denotes the value of individual $i+1$ on the j th objective function, and f_j^{i-1} denotes the value of individual $i-1$ on the j th objective function. Individuals with larger crowding degree distance are better under the same non-dominated rank sequence.

4.2.4 Genetic Operator

1. Selection

Every time, two people are chosen from the population using the binary tournament selection method, and their ranks are compared based on the Pareto in the fast non-dominated ranking. Individuals with the greater crowding degree distance are preferred until the new population size equals the initial population size if Pareto has the same rank.

2. Crossover Operator

In this paper, two-point crossover is adopted, as shown in Figure 6. Two gene fragments with the same chromosome are selected, and no other categories of goods are stored in the space between them. If the conditions are not met, the crossover operation will be abandoned.

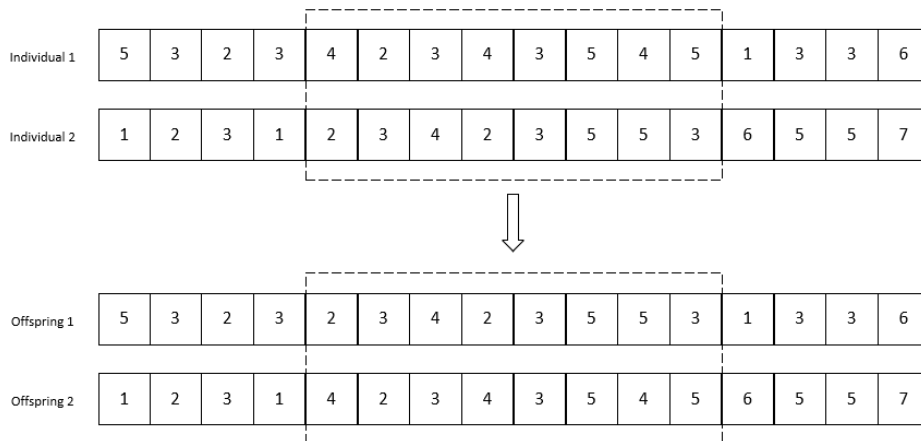


Figure 6: Crossover operation.

3. Mutation Operator

As illustrated in Figure 7, the genes of two distinct commodities are chosen at random to switch positions in order to boost the population's diversity.

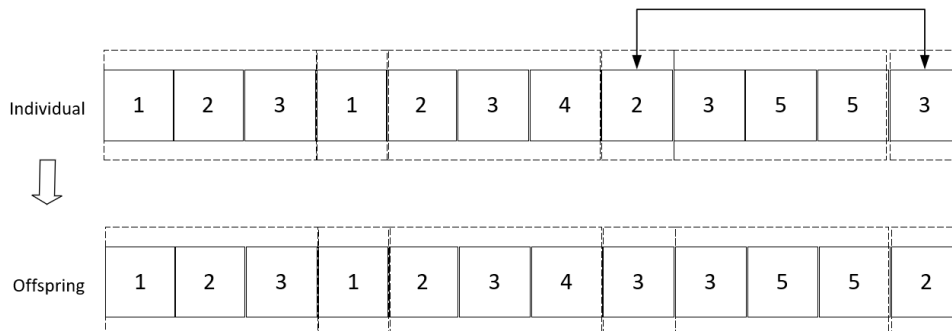


Figure 7: Mutation operations.

4.2.5 Learning Mechanisms

NSGA-II algorithm has some blindness in the process of randomly generating chromosomes as well as crossover and mutation, and most of the time, it will fall into the local optimal solution that cannot be jumped out, and NSGA-II algorithm needs to carry out the nondominated sorting, which is a more time-consuming computational step, so the efficiency of the algorithm will be reduced for large-scale problems. To solve this problem, this paper introduces a learning mechanism to improve NSGA-II on the basis of Yakun Wang et al. and Gray Wolf algorithm. The learning mechanism refers to that the progeny learn the best chromosome genes of the previous generation, improve the individual yield of the progeny, and inherit the information of the optimal solution, which can accelerate the speed of the algorithm to obtain the global optimal solution, as shown in Figure 8.

The learning mechanism is implemented as follows:

1. Two positions were randomly selected from the range of optimal solution chromosome lengths;
2. The gene segment between the two positions is copied directly to the corresponding chromosomal position in the offspring;
3. The goods category information of the remaining left gene segment is disrupted and assigned to the left gene segment of the offspring, and the goods category information of the remaining right gene segment is disrupted and assigned to the right gene segment of the offspring to generate a new offspring chromosome.

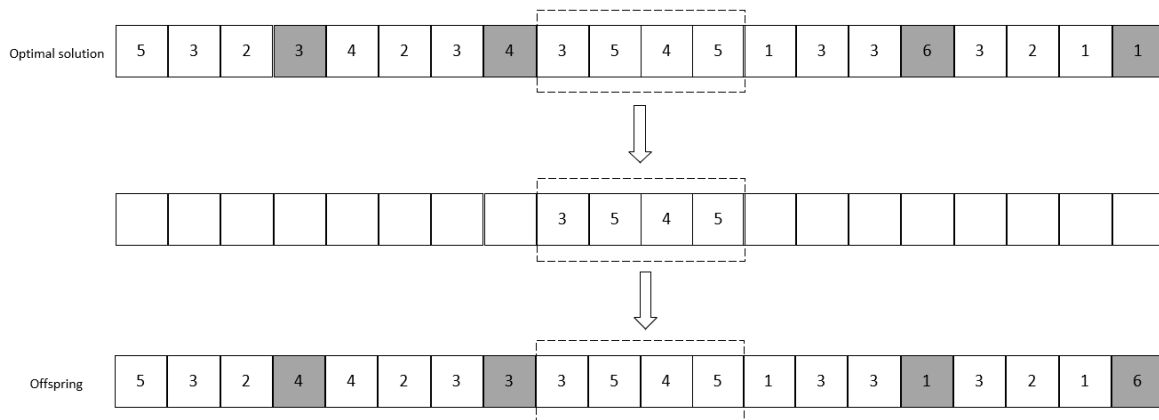


Figure 8: Basic steps of the learning mechanism.

4.2.6 Principles of Termination

Setting appropriate evolutionary termination conditions for the algorithm is essential for the realization of the optimal solution of multi-objective optimization problems because the termination principle of the algorithm directly impacts the algorithm's solution accuracy and the time it takes to obtain the optimal solution of the objective function. In this paper, the improved NSGA-II algorithm for the storage location assignment issue adopts an appropriate maximum number of iterations as its termination condition.

4.3 Chapter Conclusion

In this paper, the improved NSGA-II algorithm is used to solve the storage location assignment model. The improved algorithm inherits the advantages of the original algorithm and optimizes the shortcomings of the original algorithm to a certain extent. To be specific, the improved NSGA-II algorithm not only optimizes the blindness of the original algorithm

when crossing and mutating, but also helps the original algorithm get rid of the dilemma of being trapped in the local optimal solution. However, it is inevitable that the improved algorithm still has some shortcomings. For example, the running time of the algorithm is still not short, and the solving efficiency has not been greatly improved. This is also one of the key research points in the future.

5 Instance Validation

5.1 Description of the Example

Through field research, this paper obtains the order data set of A warehouse of H e-commerce enterprise in May 2024. This includes data such as order number, item type, initial location of the item, etc. Because the business data needs to be treated confidentially, order numbers and product categories are replaced by numbers and the original order data set was processed and cleaned to retain 250 valid orders. A warehouse contains 30 shelves, each shelf is divided into 5 layers, each layer has 6 Spaces, storing a total of 60 kinds of goods, after the warehouse runs for a period of time, some of the warehouse space is empty, that is, there is no commodities stored in the storage space, need to be replenished and put new goods on the shelf, the initial status of the shelves is shown in Table 3. Currently, the warehouse has received 250 orders that need to be selected, and each order contains no more than 10 kinds of goods. The orders to be selected are shown in Table 4. According to the order to be selected, out of the 50 types of goods that are now kept in the warehouse, some need to be restocked, and new items need to be put on the shelves. Table 5 lists the kinds of commodities that need to be restocked, along with the new items and the quantity of storage space needed. When taking into account the separate picking of every order, it is necessary to allocate all the goods to be replenished and new goods on the shelf, and to achieve the maximum efficiency of total order picking.

Table 3: Initial state of shelves.

Shelf existing product categories	Location (row, column, layer)
36	(1, 1, 3)
19	(7, 1, 2)
7	(11, 1, 1)
31	(15, 1, 4)
34	(18, 1, 2)
19	(23, 1, 2)
36	(6, 2, 5)
30	(9, 2, 5)
33	(14, 2, 2)
35	(21, 3, 1)
36	(30, 3, 3)
20	(3, 4, 1)
29	(17, 4, 5)
13	(26, 4, 2)
11	(4, 5, 2)

Table 4: Information on 250 orders to be picked.

Order Number	Type of goods included in the order (number)
1	9, 11, 13, 24, 33
2	22, 25, 29, 34, 41
3	34, 35, 36, 42, 43, 44, 59
4	9, 10, 37, 57, 58
5	11, 12, 13, 14, 15, 17, 18, 20, 44, 45
6	11, 12, 14, 15, 19, 20, 21, 45, 46
7	1, 3, 4, 6, 10, 48, 57
8	56
9	13, 15, 37, 49, 51, 55, 56, 60
10	15, 24, 52, 54
...	...
250	4, 7, 34

Table 5: Space required for goods to be replenished and new goods.

Commodity i	Number of storage spaces required to store commodity i
1	24
2	8
3	13
4	10
5	14
6	17
7	27
8	14
9	23
10	19
...	...
60	8

5.2 Example Results

Firstly, the degree of correlation between 50 existing commodities in the warehouse is calculated based on the historical orders, while the degree of correlation between all commodities is calculated based on the orders to be picked, and this correlation matrix is cumulatively added to the historical order correlation matrix, and the final results of the commodity correlation calculation are shown in Table 6.

Table 6: Results of commodity relevance calculation.

Commodity	1	2	3	4	5	6	...	59	60
1	1.000	0.190	0.109	0.146	0.109	0.111	...	0.022	0.042
2	0.190	1.000	0.132	0.146	0.229	0.050	...	0.026	0.024
3	0.109	0.132	1.000	0.116	0.100	0.103	...	0.026	0.024
4	0.146	0.146	0.116	1.000	0.143	0.119	...	0.023	0.022
5	0.109	0.229	0.100	0.143	1.000	0.132	...	0.053	0.024
6	0.111	0.050	0.103	0.119	0.132	1.000	...	0.054	0.024
...
59	0.022	0.026	0.026	0.023	0.053	0.054	...	1.000	0.258
60	0.042	0.024	0.024	0.022	0.024	0.024	...	0.258	1.000

Based on the algorithm derived from the enhanced NSGA-II design and the storage site assignment model outlined in the preceding section, the initial and improved storage space method allocation results are obtained under the Matlab R2023(b) programming environment, which are shown in Figure 9 and Figure 10, and the algorithm convergence curves are shown in Figure 11.

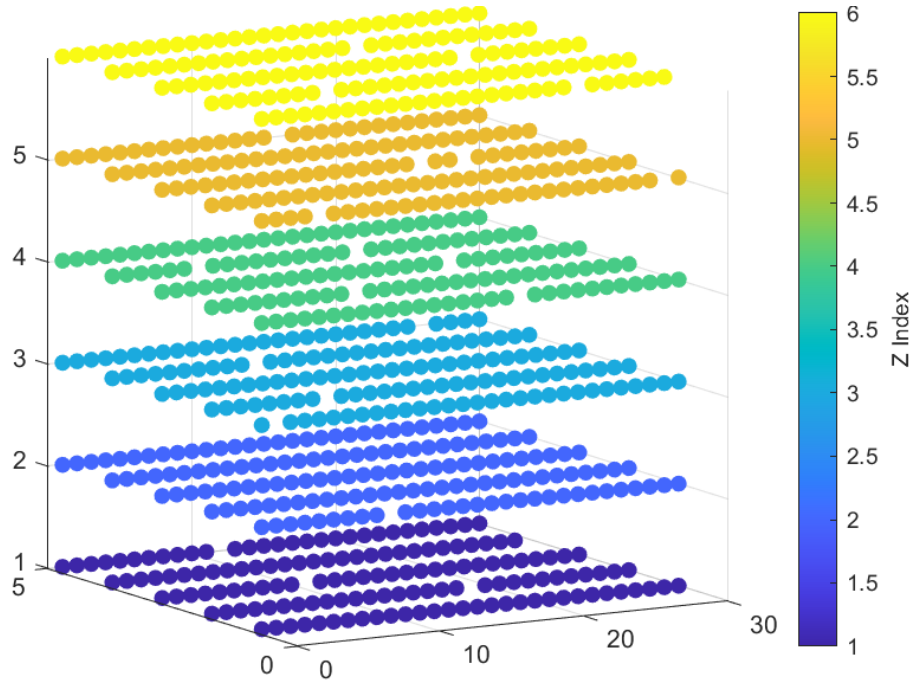


Figure 9: Results of storage space allocation before improvement.

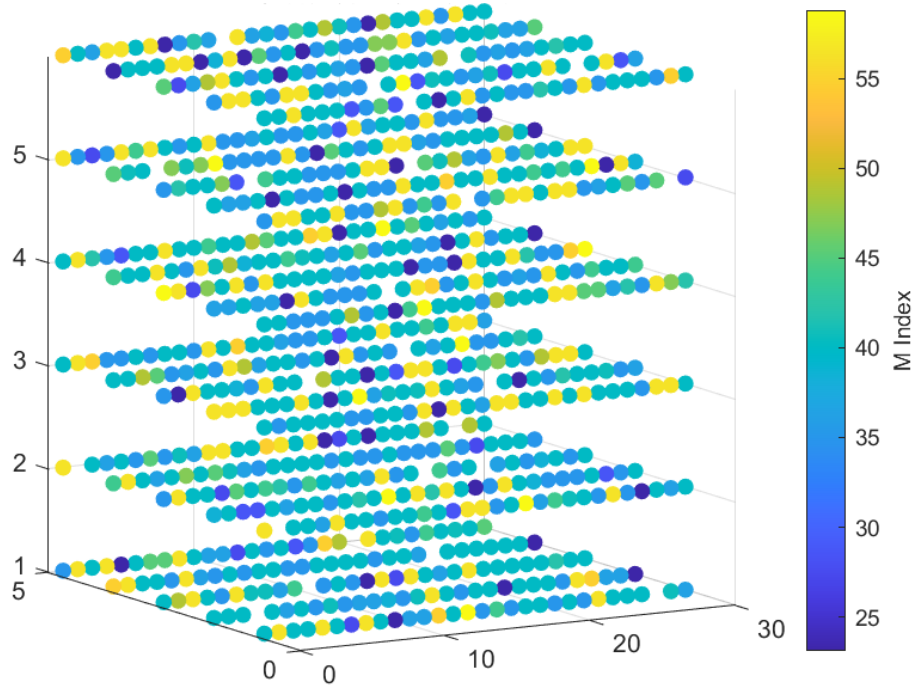


Figure 10: Results of storage space allocation obtained based on improved NSGA-II algorithm.

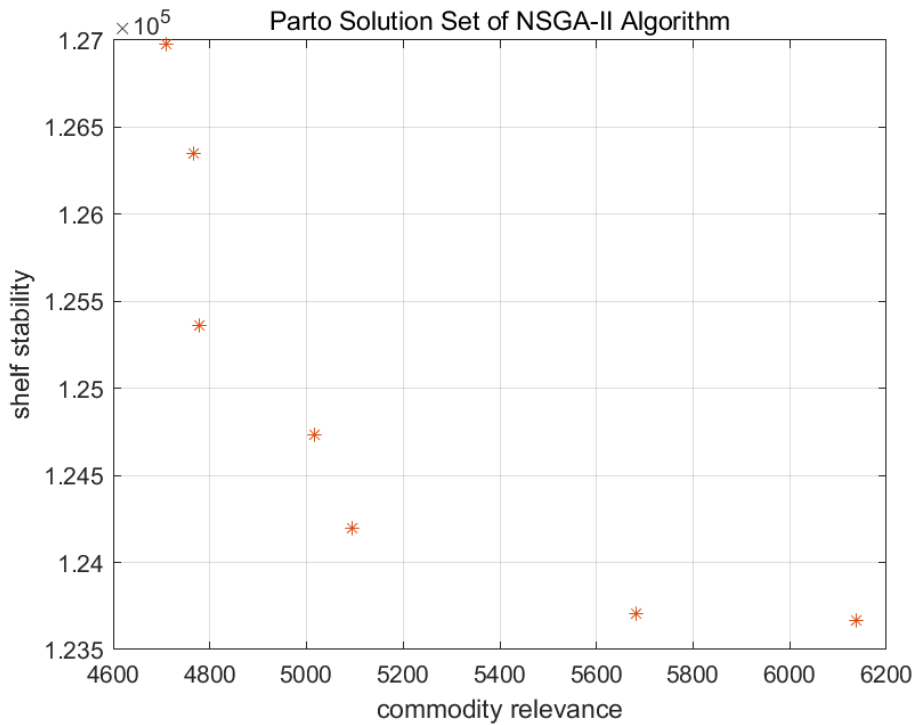


Figure 11: Pareto optimal curve based on improved NSGA-II algorithm.

In the process of completing the selection of 250 orders, this paper makes a detailed comparison between the effect of space allocation before and after improvement as shown in Table 7. It can be seen from the table that optimal objective function value 1 and optimal objective function value 2 obtained by the original algorithm are 6373.3 and 127350 respectively. However, the optimal objective function value 1 and the optimal objective function value 2 obtained by the improved algorithm are 4802.7 and 126670 respectively. The optimization percentages were 32.7% and 0.54%, respectively. This shows that the improved algorithm and the shelf allocation strategy based on the correlation degree of quantity and distance of goods proposed in this paper can achieve significant optimization of the objective function.

Table 7: Comparison of results of storage space allocation strategies before and after improvements.

	Optimal objective function value I	Optimal objective function value II
Pre-improvement	6373.3	127350.0
Post-improvement	4802.7	126670.0
Percentage optimization	32.70%	0.54%

In addition, since the change trend of fitness (objective function value) reflects the optimization process of the algorithm, this paper will further verify the effectiveness of the improved algorithm through the fitness convergence curve. The fitness convergence curves of the original NSGA-II algorithm and the improved NSGA-II algorithm are shown in Figure 12. It can be seen from the figure that the initial position of the improved algorithm is optimized and the convergence speed is faster. It not only increases the diversity of the population, but also makes the algorithm jump out of the local extreme value, and the calculation result is closer to the global optimal.

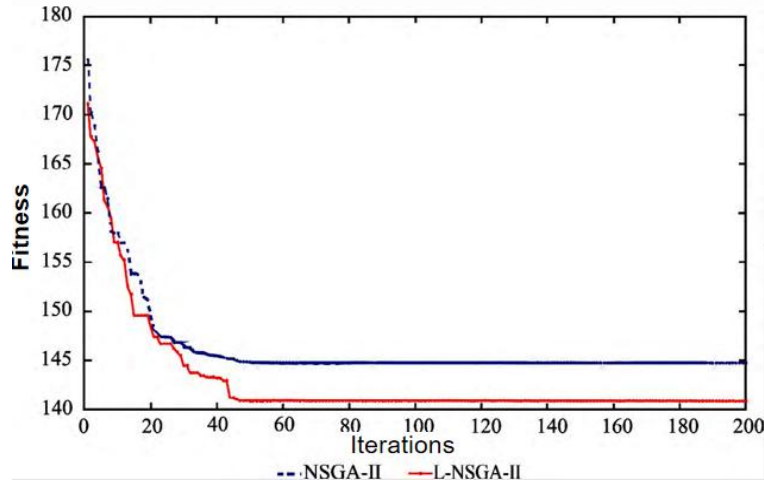


Figure 12: The fitness convergence curves of the NSGA-II algorithm and the improved algorithm.

5.3 Model Sensitivity Analysis

In this study, we do a sensitivity analysis of shelf specifications to examine how modifications to them affect order choosing and storage space allocation. Warehouses can flexibly set the number of storage spaces in each level of a shelf by adding or subtracting partitions to change the shelf specifications and accommodate different types of goods. Accordingly, the more storage spaces there are, the smaller the storage capacity becomes. It is very important to examine how the amount of rack storage spaces affects the storage space allocation scheme since it influences the degree of commodity dispersion, which in turn influences the effectiveness of order picking.

Figure 13 shows that the increase in the number of storage spaces can effectively reduce the objective function value. In-depth analysis can be found that the increase in the number of sub-bin can enable the goods to fully realize decentralized storage to match the order embodied in the category relationship network, so that order picking can be accomplished by moving fewer times of shelves. Furthermore, the objective function's value exhibits a varying downward trend as the number of storage spaces increases. It is evident that the impact of expanding storage capacity on choosing efficiency improvement is progressively waning.

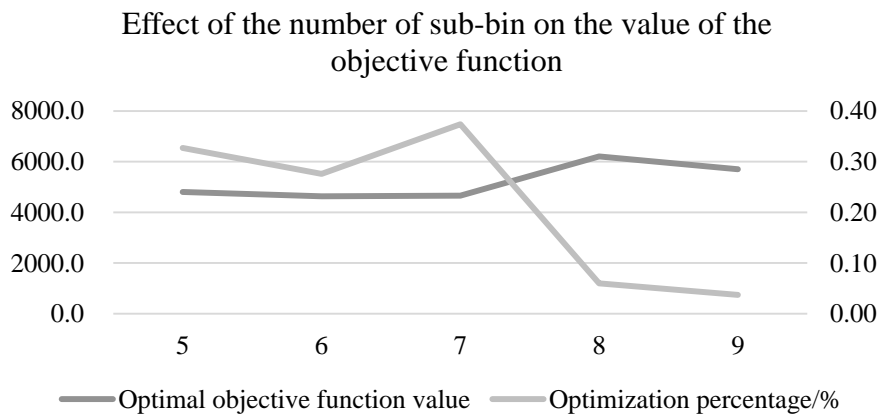


Figure 13: Effect of the amount of storage space in each layer on the value of the objective function.

The order type is another important factor affecting commodity storage location assignment, such as the beauty warehouse order has the characteristics of "one single multi-product" and the electronic warehouse orders have the characteristics of "small batch, multi-frequency". In order to further verify the effect of the order type, the order is divided into three cases containing the product types in the range [1,3], [1,5] and [1,7]. At the same time, the fitness (objective function value) of the three cases in different storage spaces are solved respectively. The analysis results are shown in Figure 14.

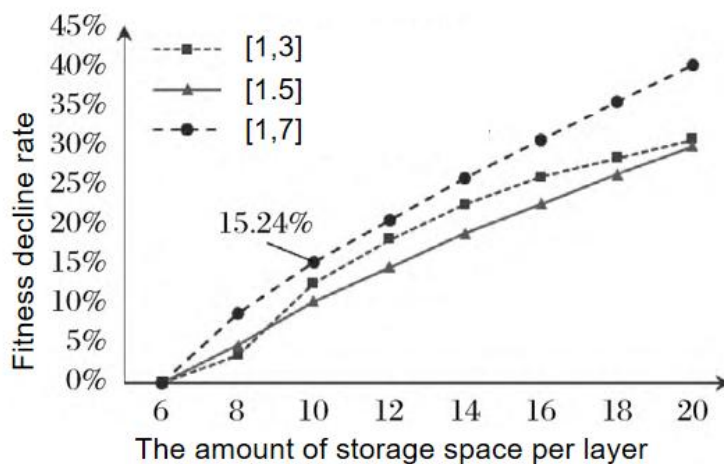


Figure 14: Effect of the amount of storage space in each layer on the value of the objective function under different order types.

Figure 14 shows that the effect of the amount of storage space in each layer on reducing the value of the objective function gradually weakens and becomes stable. The fewer types of goods included in the order, the faster the effect decreases. Increasing the amount of storage space can reduce the number of shelf movements, however, reducing the shelf efficiency of goods in the stocking stage may produce more stocking costs, so the warehouse manager should consider the cost of picking and loading goods, and set the appropriate amount of storage space for different order types.

6 Conclusions

In recent years, with the rapid growth of e-commerce, in order to improve the efficiency of order picking and distribution, e-commerce enterprises have begun to gradually implement the "goods to people" picking system. In this paper, a method of storage location assignment based on commodity correlation degree is proposed which considering the quantity of goods - distance correlation, and a mathematical model is constructed with the objective of optimizing the correlation degree between commodities and the stability of shelves. The improved NSGA-II algorithm is designed for solving the problem, and the example verification proves that the space allocation model and algorithm proposed in the paper can significantly optimize the value of the objective function, and provide corresponding management suggestions for the effective operation of the movable shelf storage system. Specifically, the model and algorithm proposed in this study are applied to solve the problem of commodity storage allocation. It can effectively reduce the handling frequency of shelves, improve the efficiency of picking, and lay a sound scheduling foundation for the sustainable development of e-commerce enterprises. In addition, for picking systems with increasingly complex order types, administrators can take full advantage of the characteristics of decentralized storage, increase the number of shelf

storage Spaces to reduce the number of shelf movements, thereby reducing warehouse energy consumption (such as electricity and transportation fuel), and reduce carbon emissions to help enterprises achieve green development goals.

The influence of the correlation between the goods and the shelf's stability on the space allocation method is taken into consideration in this work when examining the storage site assignment problem of the "goods to person" choosing system, assuming that the shelf's initial state is not empty. However, in the actual operation process, there are other factors that can affect the results of space allocation. This paper fails to incorporate all the influencing factors into the model, which is not fully practical. Therefore, in future studies, we will consider adding more influencing factors, such as the frequency of all goods on the shelf out of stock and social factors such as carbon emissions. In addition, the improved NSGA-II algorithm proposed in this paper is still a little insufficient in computing efficiency, which is also an important research content expected to be improved in the future.

Data availability

The data sets that were created and/or analyzed as part of the current study are not publicly available due to proprietary reasons, but can be obtained from the corresponding author on request.

Funding

The authors would like to appreciate the reviewers for all helpful comments, and to thank the foundation of National Social Science Foundation Annual Project (19BGL097).

About the Authors

Xinqi Wu (2000-), female, born in Yancheng, Jiangsu Province, master's candidate, research direction is logistics and supply chain management

Juping Shao (1971-), male, born in Zhuanglang, Gansu Province, professor, doctor, master tutor, research direction is logistics and supply chain management.

Yanan Sun (1971-), male, born in Guannan, Jiangsu Province, Professor of industry in Jiangsu Province, master tutor, President of Suzhou Youlesai Supply Chain Management Co., LTD., research direction is logistics technology and equipment

Authors and Affiliations

Xinqi Wu, Juping Shao and Yanan Sun have contributed equally to this work.

School of Business, Suzhou University of Science and Technology, Suzhou 215009, China

Anwood Logistics System Co. Ltd, Suzhou 215021, China

Contributions

X.W. was responsible for building the model and writing the manuscript. J.S. managed the data, interpreted it, and reviewed and approved the final manuscript. Y.S. processes the model

algorithm accordingly.

Open Access

I confirm that I understand Energy, Sustainability and Society is an open access journal that levies an article processing charge per articles accepted for publication. By submitting my article I agree to pay this charge in full if my article is accepted for publication.

Competing Interests

No, I declare that the authors have no competing interests as defined by BMC, or other interests that might be perceived to influence the results and/or discussion reported in this paper.

Third Party Material

All of the material is owned by the authors and/or no permissions are required.

References

- [1] *The United Nations Conference on Trade and Development (UNCTAD) Home Page*. [cited 2024 10 October]; Available from: <https://unctad.org/>.
- [2] Wisittipanich, W. and C. Kasemset, *Metaheuristics for warehouse storage location assignment problems*. Journal of CMUJ NS Special Issue on Logistics and Supply Chain Systems, 2015. **14**(4).
- [3] Cordeau, J.-F., P. Legato, and R.M. Mazza, *Integrating storage allocation with manual order picking and replenishment operations in a distribution centre*. International Journal of Production Research, 2024: p. 1-17.
- [4] Peterson, T. and M.G. Gonzalez, *Quantifying Warehouse Automation and Sustainability*. 2022.
- [5] Hong, S., *Order batch formations for less picker blocking in a narrow-aisle picking system*. Industrial Engineering and Management Systems, 2015. **14**(3): p. 289-298.
- [6] Koch, S. and G. Wäscher, *A grouping genetic algorithm for the order batching problem in distribution warehouses*. Journal of Business Economics, 2016. **86**: p. 131-153.
- [7] Jiang, Z.-Z., et al., *Spatial and temporal optimization for smart warehouses with fast turnover*. Computers & Operations Research, 2021. **125**: p. 105091.
- [8] Ardjmand, E., et al., *Minimizing order picking makespan with multiple pickers in a wave picking warehouse*. International Journal of Production Economics, 2018. **206**: p. 169-183.
- [9] Xie, J., et al., *A bi-level optimization model for grouping constrained storage location assignment problems*. IEEE transactions on cybernetics, 2016. **48**(1): p. 385-398.

- [10] Hanson, R., L. Medbo, and M.I. Johansson, *Performance characteristics of robotic mobile fulfilment systems in order picking applications*. IFAC-PapersOnLine, 2018. **51**(11): p. 1493-1498.
- [11] De Koster, R., T. Le-Duc, and K.J. Roodbergen, *Design and control of warehouse order picking: A literature review*. European journal of operational research, 2007. **182**(2): p. 481-501.
- [12] Xiao, J. and L. Zheng, *A correlated storage location assignment problem in a single-block-multi-aisles warehouse considering BOM information*. International Journal of Production Research, 2010. **48**(5): p. 1321-1338.
- [13] Pang, K.-W. and H.-L. Chan, *Data mining-based algorithm for storage location assignment in a randomised warehouse*. International Journal of Production Research, 2017. **55**(14): p. 4035-4052.
- [14] Jin, X. and Y. Li. *Research on Storage Location Assignment of B2C Electronic Commerce Enterprise Based on Association Rules Mining*. in *LISS 2013: Proceedings of 3rd International Conference on Logistics, Informatics and Service Science*. 2015. Springer.
- [15] Brynzér, H. and M.I. Johansson, *Storage location assignment: Using the product structure to reduce order picking times*. International Journal of Production Economics, 1996. **46**: p. 595-603.
- [16] Zhou, G. and L. Mao, *Design and simulation of storage location optimization module in AS/RS based on FLEXSIM*. International Journal of Intelligent Systems and Applications, 2010. **2**(2): p. 33.
- [17] Heskett, J.L., *Cube-per-order index-a key to warehouse stock location*. Transportation and distribution Management, 1963. **3**(1): p. 27-31.
- [18] Yang, C.-L. and T.P.Q. Nguyen, *Constrained clustering method for class-based storage location assignment in warehouse*. Industrial Management & Data Systems, 2016. **116**(4): p. 667-689.
- [19] Hausman, W.H., L.B. Schwarz, and S.C. Graves, *Optimal storage assignment in automatic warehousing systems*. Management science, 1976. **22**(6): p. 629-638.
- [20] Yu, Y. and R.B. De Koster, *Optimal zone boundaries for two-class-based compact three-dimensional automated storage and retrieval systems*. IIE transactions, 2009. **41**(3): p. 194-208.
- [21] Mirzaei, M., N. Zaerpour, and R.B. de Koster, *How to benefit from order data: correlated dispersed storage assignment in robotic warehouses*. International Journal of Production Research, 2022. **60**(2): p. 549-568.
- [22] Zhen-ping, L., F. Xin-ran, and W. Ling-yun, *Study on the storage allocation problem under cargo to person picking mode*. Operations Research and Management Science, 2020. **29**(2): p. 1.

- [23] Frazele, E. and G.P. Sharp, *Correlated assignment strategy can improve any order-picking operation*. Industrial Engineering, 1989. **21**(4): p. 33-37.
- [24] Xiang, X., C. Liu, and L. Miao, *Storage assignment and order batching problem in Kiva mobile fulfillment system*. Engineering Optimization, 2018. **50**(11): p. 1941-1962.
- [25] Guan, M. and Z. Li, *Genetic Algorithm for scattered storage assignment in Kiva mobile fulfillment system*. American Journal of Operations Research, 2018. **8**(6): p. 474-485.
- [26] Muppani, V.R. and G.K. Adil, *Efficient formation of storage classes for warehouse storage location assignment: a simulated annealing approach*. Omega, 2008. **36**(4): p. 609-618.
- [27] Yang, P., et al., *An integrated optimization of location assignment and storage/retrieval scheduling in multi-shuttle automated storage/retrieval systems*. Journal of Intelligent Manufacturing, 2015. **26**: p. 1145-1159.
- [28] Hu X P, Ding T R, Zhang Y K, et al. A method for storage assignment in robotic mobile fulfillment system based on association networks[J]. Frontiers of Science and Technology of Engineering Management, 2022, 41(1): 56-64.
- [29] Shi, X., et al., *A bi-level optimization approach for joint rack sequencing and storage assignment in robotic mobile fulfillment systems*. Science China Information Sciences, 2023. **66**(11): p. 212202.
- [30] Kim, H.-J., C. Pais, and Z.-J.M. Shen, *Item assignment problem in a robotic mobile fulfillment system*. IEEE Transactions on Automation Science and Engineering, 2020. **17**(4): p. 1854-1867.
- [31] Ding T R, Zhang Y K, Wang Y Y, et al. An item allocation method for “manual-robotic” dual picking systems based on association networks[J]. Chinese Journal of Management Science, 2023, 31(3): 26-37.
- [32] Wang Z, Lu P, Hu X P. Method of locating racks in mobile-rack warehousing system[J]. Chinese Journal of Management Science, 2024, 32(8): 84-94.
- [33] Zhou, L., et al., *Stochastic models of routing strategies under the class-based storage policy in fishbone layout warehouses*. Scientific Reports, 2022. **12**(1): p. 12876.
- [34] Deb, K., et al., *A fast and elitist multiobjective genetic algorithm: NSGA-II*. IEEE transactions on evolutionary computation, 2002. **6**(2): p. 182-197.
- [35] Gao, K. and L. Xu, *A new prediction-based evolutionary dynamic multiobjective optimization algorithm aided by Pareto optimal solution estimation strategy*. Applied Soft Computing, 2024. **165**: p. 112022.
- [36] Uma V, Lin E, Jayashree D. Efficient storage location assignment using genetic algorithm in warehouse management system[J]. Journal of Advanced Research, 2016, 4(6): 18-24.

- [37] Chen L, Langevin A, Riopel D. The storage location assignment and interleaving problem in an automated storage/retrieval system with shared storage[J]. *International Journal of Production Research*, 2010, 48(4): 991-1011.
- [38] Zhai M Y, Wang Z, Li Y T, et al. Research on the problem of allocating commodity storage locations in a mobile-rack storage system[J]. *Chinese Journal of Management Science*, 2023, 31(3): 167-176.