Research on Workshop Layout Based on Genetic Algorithm of Machine Learning K-means Clustering

Hongrun Pang*[1], Chengjun Ji[2]

School of Business Administration, Liaoning Technical University, Huludao 125000, PR China

E-mail: 1368497524@qq.com*[1]; jichengjun@lntu.edu.cn[2]

Abstract. Aiming at the problems of logistics confusion and low efficiency among equipment caused by unreasonable workshop layout, this problem can be effectively solved by optimizing mathematical model and adopting improved genetic algorithm based on multi-objective. On the basis of classical genetic algorithm, a multi-strategy parallel genetic algorithm based on machine learning is proposed, and the performance of genetic algorithm is improved by using machine learning method. Firstly, the parallel idea is used to accelerate the evolution process of genetic algorithm, and the initial population is divided into multiple clusters by using K-means clustering algorithm. Then, reinforcement learning is introduced to realize the self-learning of the crossover probability of important parameters in genetic algorithm, so that the crossover probability can adapt to the evolution process according to experience. The experimental results show that the multi-strategy parallel genetic algorithm of machine learning is obviously superior to the classical genetic algorithm, which can optimize the original layout of the workshop well and improve the effect significantly.

Keywords: Workshop layout; genetic algorithm; machine learning; K-means clustering

1 Introduction

With the change of the market, in order to cater to the market and continuously reduce the production cost, enterprises improve the production process of products and optimize the layout of workshops reasonably[1]. The layout of workshop equipment is still one of the most complicated problems in intelligent production at present, and its essence is a combinatorial optimization problem, which can reasonably arrange all units in the workshop to improve the utilization rate of workshop space and reduce the logistics cost in the workshop to reduce the production cost of enterprises. Therefore, the rationality of workshop equipment layout is particularly critical to the production system of enterprises.

In recent years, the research on workshop layout optimization mostly focuses on intelligent optimization algorithms such as genetic algorithm [2], simulated annealing algorithm [3] , tabu search algorithm[4], artificial fish swarm algorithm[5].and particle swarm optimization algorithm[6] .Lu Yizhen [7]and others introduced adaptive genetic

T. Yao et al. (eds.), *Proceedings of the 2024 3rd International Conference on Engineering Management and Information Science (EMIS 2024)*, Advances in Computer Science Research 111,
https://doi.org/10.2991/978-94-6463-447-1_20

operator strategy to improve the genetic operator in the process of solving the model of workshop layout problem by genetic simulated annealing algorithm; Liu Xiaopeng [8] and others studied an optimization algorithm based on decomposition and coordination to solve the problem that the system performance index of multi-layer flow shop does not have a closed mathematical expression, thus providing an important decision-making basis for solving the workshop layout planning problem; Zhang Siqi [9]and others adopted the multi-objective non-dominated migratory bird optimization (NSMBO) algorithm to solve the workshop facility optimization problem, making the migratory bird algorithm more suitable for solving the workshop facility layout problem.

In the optimization of workshop layout facilities, most of the logistics costs between operating units are taken as the optimization goal, but the non-logistics closeness between operating units is not taken into account [10]. In this paper, considering the different non-logistics situations among different operating units, aiming at the minimum logistics cost and the optimal non-logistics closeness, a genetic algorithm based on machine learning K-means clustering is introduced [11]. Firstly, the initial population is clustered by K-means clustering to ensure the diversity of the initial population, and at the same time, reinforcement learning that can perceive the environment independently is introduced to realize the self-learning of the crossover probability of important parameters in the genetic algorithm. The crossover probability is adapted to the evolution process according to experience, and various strategies are cited to improve the performance of genetic algorithm. The optimization model of workshop facilities layout is solved, and the effectiveness of the model and algorithm is verified by an example analysis.

2 Optimization Model of Workshop Facilities Layout

2.1 Model Hypothesis

Workshop equipment layout can be divided into four types according to equipment arrangement: one-way linear distribution, multi-line linear distribution, U-shaped distribution and circular distribution. This paper takes the distribution of multi-line linear equipment in the workshop as the research object, and the method to solve this problem is to arrange the equipment positions in the workshop reasonably. In order to facilitate the study, simplify the equipment layout of the workshop and make the following assumptions:

1) Because the height of the general workshop building is far greater than the height of the equipment, only the two-dimensional plane layout of the workshop is considered;

2) All the units in the workshop are in the same plane, with the lower left corner as the origin of the plane;

3) Each unit shape of workshop equipment is a rectangle with known length and width, and the length and width are distributed parallel to the X axis and Y axis of the coordinate system.

In the coordinate system XOY, L and W represent the length and width of the workshop respectively; L_j represents the length of device j; W_j represents the width of device j.

2.2 Multi-Objective Optimization Function

The purpose of equipment layout optimization in production workshop is to plan a reasonable logistics transportation path, reduce the production cost of enterprises, and improve the continuity of equipment logistics. For this purpose, this paper establishes a multi-objective optimization function with the minimum total cost of material handling between equipment and the maximum comprehensive correlation (non-logistics closeness) as the optimization goal.

Material handling cost C_1 is:

$$minC_1 = \sum_{i=1}^{n}\sum_{j=1}^{n} P_{ij}D_{ij}Q_{ij} \tag{1}$$

Where: C_1 represents the total cost of material handling in the production workshop, P_{ij} represents the unit cost of material handling between equipment I and equipment J, D_{ij} represents the logistics distance between equipment I and equipment J, and Q_{ij} represents the material flow (quantity) between equipment I and equipment J.

The function $C_2$of non-logistics relationship-comprehensive correlation (non-logistics closeness) is:

$$maxC_2 = \sum_{i=1}^{n}\sum_{j=1}^{n} v_{ij}b_{ij} \tag{2}$$

Where v_{ij} represents the strength of non-logistics relationship between equipment I and equipment, and b_{ij}represents the distance proximity between equipment I and equipment J.

The non-logistics relationship between equipment also plays a great role in the smooth operation of the workshop, so it is very important to analyze the closeness of the non-logistics relationship between equipment. This paper analyzes the closeness of the non-logistics relationship between equipment through five aspects (process continuity, commonality of auxiliary tools, convenience of management and communication, environmental safety and rationality of energy transmission) . The closeness between devices is divided into six grades and the corresponding values of each grade are shown in Table 1.

Table 1. Classification table of non-logistics closeness between equipment

Correlation grade	Intimacy	Value of v_{ij}
A	Necessity to approach	4
E	Particularly important	3
I	Important	2
O	Generally close	1
U	Unimportance	0
X	Not close	-1

The proximity degree is determined by the ratio of the distance D_{ij}between equipment I and equipment J and the maximum distance D_{max} between workshop equipment. It is also divided into 6 grades, and the adjacent comparison table is shown in Table 2.

Table 2. Proximity comparison table

Value range of l_{ij}	The corresponding b_{ij} factor is selected
$0 < D_{ij} \le D_{max}/6$	1
$D_{max}/6 < D_{ij} \le D_{max}/3$	0.8
$D_{max}/3 < D_{ij} \le D_{max}/2$	0.6
$D_{max}/2 < D_{ij} \le 2D_{max}/3$	0.4
$2D_{max}/3 < D_{ij} \le 5D_{max}/6$	0.2
$5D_{max}/6 < D_{ij} \le D_{max}$	0

WhereD_{max} represents the maximum distance between two working units, that is, the sum of the length and width of the workshop.

Simplify the mathematical model, transform the double-objective model into a single-objective model, and get the objective function F.

$$minF = w_1 u_1 \sum_{i=1}^{n}\sum_{j=1}^{n} P_{ij} D_{ij} Q_{ij} - w_2 u_2 \sum_{i=1}^{n}\sum_{j=1}^{n} v_{ij} b_{ij} \tag{3}$$

Where w_1 represents the weight of logistics handling cost, and w_2 represents the weight of non-logistics relationship (comprehensive correlation), and the sum of the two weights is 1.

Where u_1、u_2 represent normalization factors, and their functional expressions are as follows:

$$u_1 = \frac{1}{\sum_{i=1}^{n}\sum_{j=1}^{n} P_{ij} D_{max} Q_{ij}};\ \ u_2 = \frac{1}{\sum_{i=1}^{n}\sum_{j=1}^{n} v_{ij}}. \tag{4}$$

2.3 Constraint Condition

There are various constraints in the layout of workshop equipment, including the constraints of process route, manual operation space and equipment appearance, so the constraints of the equipment layout model in this paper are as follows:

1)Each piece of equipment should be in the workshop, that is:

$$L_0 < x_i - \frac{L_i}{2} < L - L_0; W_0 < y_i - \frac{W_i}{2} < W - W_0 \tag{5}$$

Where L_0 represents the minimum distance between the equipment and the boundary in the transverse direction; W_0 indicates the minimum distance between the equipment and the boundary in the longitudinal direction.

2)The horizontal direction shall not exceed the scope of the workshop：

$$\sum_{i=1}^{n}\sum_{j=1}^{n}(|x_i - x_v| + \frac{L_i + L_v}{2}) \le L \tag{6}$$

3)The vertical direction shall not exceed the scope of the workshop:

$$\sum_{i=1}^{n}\sum_{j=1}^{n}(|y_i - y_j| + \frac{W_i + W_j}{2}) \leq W \tag{7}$$

4)Each device can only appear once, that is:

$$\sum_{i=1}^{n} Z_i = n \tag{8}$$

Where: $Z_i = 1$ means that this equipment appears once.

3 K-Means Clustering Genetic Algorithm for Machine Learning

3.1 Algorithm Principle Description

In this paper, two strategies are cited to improve the classical genetic algorithm, and a machine learning K-means clustering (MLGA) multi-strategy parallel genetic algorithm is designed to solve this model. First of all, in order to ensure the evolution process of the algorithm, the K-means clustering method is applied to divide the initial population of the genetic algorithm, maintain the diversity and uniformity of the population, and make the sub-populations communicate during the whole evolution process. On this basis, the machine learning method is used to adaptively adjust the crossover probability in the genetic algorithm, so that the crossover probability can be adaptively evolved through experience.

3.2 Algorithm Design

3.2.1 Chromosome Coding and Decoding

In dealing with the workshop layout problem, genetic algorithm uses real number coding to code chromosomes, which changes the binary coding mode of traditional genetic algorithm. This method can simplify the complicated workshop equipment layout and improve efficiency. If the total length of the arranged equipment exceeds the total length of the workshop, the next line will be automatically arranged for equipment arrangement.

3.2.2 Population Initialization and Sub-Population Division

First, a random population will be initialized with the scale of n, and then the sub-population will be divided by the strategy of K-means clustering.

The strategy method of K-means clustering designed in this paper divides the individuals after the initial population, evenly divides the same kind of individuals, and obtains a number of sub-populations with large internal differences, thus increasing the dispersion of sub-populations in space and enhancing the dispersion and global optimization ability of the algorithm.

3.2.3 Fitness Evaluation and Selection

Fitness function is a standard function to judge the advantages and disadvantages of sub-populations. According to the objective function (3) of this paper is to solve the minimization model, the fitness function of this paper can be determined as follows:

$$f(F) = \frac{1}{F} \tag{9}$$

Choose. Use roulette and the method of combining with elite retention strategy. The probability that the individual I is retained is:

$$p(i) = \frac{f_i}{\sum f_i(F)} \tag{10}$$

The stronger the individual's adaptability, the greater the probability of being retained.

3.2.4 Crossover and Variation

Cross. The region between the two intersections is defined as the matching crossover region, and the genes in the two parent matching crossover regions are exchanged. For the chromosome arrangement with digital coding, some codes will be repeated in simple region exchange, resulting in some codes missing. The solution to this problem is collectively called Partially Matched Crossover (PMX).In this paper, a method similar to PMX is used for crossover operation, which is suitable for integer coding.

When solving specific problems, the crossover probability parameters of genetic algorithm often need many experiments or are determined according to specific practical experience, so the parameters are closely related to the ability of genetic algorithm. The machine learning method is used to adaptively adjust the crossover probability in genetic algorithm, so that the crossover probability can evolve adaptively through experience.

In order to make the change of crossover probability parameters of genetic algorithm meet the needs of population evolution, crossover probability is introduced into genetic algorithm, and a parameter adjustment strategy based on Q-learning algorithm is designed.

The Q-learning algorithm flow is designed as follows:

(1) State space

The state space is defined as a decimal number representing the crossover probability of genetic algorithm, and the range is [0.70, 1.00).

(2) Action space

The action space is set as the change of cross probability, which is represented by {-1, 0, 1} for decreasing, unchanging and increasing, respectively, and the change range of its action is set as 0.01. Therefore, the scale of Q-Tabled is 300×3. The next crossover probability is:

$$s_{t+1} = s_t + 0.01 \times action \tag{11}$$

(3) Incentive mechanism

The reward mechanism is based on two indexes of the group, namely the average fitness and the best fitness. The reward mechanism is divided into four situations: ① the average fitness increases; ② The best fitness is increased; ③ Both increase at the same time; ④ Other circumstances. Reward and punishment are different in every situation.

(4) Greedy strategy

Through ε-greedy Green's strategy, this method can mine rules beyond the existing experience, thus preventing the premature phenomenon of this method to some extent. ε is the possibility of getting the best result from the existing experience.

(5) Process of parameter adjustment strategy

Firstly, the initial values of learning rate α, discount rate γ and discovery rate ε are set. On this basis, using reinforcement learning method, according to the current behavior, in the behavior space, by searching the maximum Q value of behavior or adopting greedy strategy, the crossover probability of current behavior is adjusted. On this basis, the Q-Table is updated and fed back, so as to realize an adaptive decision. In many scenes, after repeated training, the model can correct the parameters according to its own experience, so as to achieve the purpose of correcting the accuracy.

Variation. In this paper, the method of inversion mutation is used for mutation operation, and two mutation points appear randomly on the chromosome. The gene between the two mutation points is inverted to obtain a new chromosome.

4 Case Verification Research

4.1 Basic Situation of Workshop

This paper takes a production workshop as the background to verify the effectiveness of MLGA .The production workshop is 380m long and 190m wide, the distance between equipment and workshop boundary is not less than 8m, and the gap between equipment is at least 3m, in which the equipment area table is shown in Table 3, the comprehensive logistics correlation table is sown in Table 4, the freight table is shown in Table 5, and the handling cost table is shown in Table 6.

Table 3. Equipment area table

Number	Device name	Long	Wide
1	Warehouse blank area	100	107
2	Processing zone I	135	70
3	Repair area	36	33
4	Material storage area	204	110
5	Processing zone II	51	45
6	Polishing area	38	52
7	Processing zone III	26	33
8	Finished product area	22	28

Table 4. Comprehensive logistics correlation table

equipment	Comprehensive correlation degree of each equipment							
	1	2	3	4	5	6	7	8
1	0	O	I	I	A	U	I	U
2	O	0	I	E	E	I	U	U
3	I	I	0	I	U	U	O	U

4	I	E	I	0	E	O	U	U
5	A	E	U	E	0	I	U	U
6	U	I	U	O	I	0	U	U
7	I	U	O	U	U	U	0	U
8	U	U	U	U	U	U	U	0

Table 5. Freight table

equip-ment	Freight volume between equipment							
	1	2	3	4	5	6	7	8
1	0	0	0	0	5600	5600	0	0
2	0	0	0	2250	4670	200	0	0
3	0	0	0	100	0	0	0	0
4	0	2250	100	0	1300	0	0	0
5	5600	4670	0	1300	0	0	0	0
6	560	200	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0

Table 6. Handling fee table

equip-ment	Handling cost between equipment							
	1	2	3	4	5	6	7	8
1	0	0.08	0.1	0.085	0.07	0.07	0	0
2	0．08	0	0.08	0.08	0.09	0.09	0	0
3	0.1	0.08	0	0.09	0	0	0	0
4	0.085	0.08	0.09	0	0.08	0	0	0
5	0.07	0.09	0	0.08	0	0	0	0
6	0.09	0.09	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0

The original layout scheme of the workshop is [5, 2, 3, 6, 4, 8, 7, 1], and the workshop length and boundary distance are limited, so the original layout scheme is the first behavior equipment 5, 2, 3, 6 and the second behavior equipment 4, 8, 7, 1. The calculated material handling cost C_1 is 536897.4 yuan, and the comprehensive correlationC_2 is 47.3.In order to find the objective function F in this paper, the different weighting factors $w_1$and $w_2$are set to be 0.7 and 0.3 respectively, and the value of the objective function F is 0.03498 after calculation.

4.2 Parameter Setting and Experimental Results and Analysis

According to the above data and the original layout, Matlab software is used for simulation optimization, and the parameters of GA and MLGA refer to the parameter setting of Zhang Yun's article[11].

Run the two algorithms for many times under the above parameters to ensure the accuracy of the results and ensure no systematic errors.

The optimal scheme of the running results of the classical genetic algorithm is [5, 2, 3, 8, 7, 4, 6, 1]. The optimal solution of multi-strategy parallel genetic algorithm layout based on machine learning K-means clustering (MLGA) is [7, 5, 6, 2, 8, 1, 3, 4].

In order to facilitate the comparison of the results, the original scheme, the results of classical genetic algorithm and the results of multi-strategy parallel genetic algorithm based on machine learning K-means clustering (MLGA) are integrated and classified. The original workshop data and the results of the two algorithms are shown in Table 7, and the final result of multi-strategy parallel genetic algorithm based on machine learning K-means clustering is better than that of classical genetic algorithm.

The original scheme was also optimized, and the cost was optimized by 13.41% and 29.74% respectively, and the comprehensive correlation was optimized by 1.27% and 6.34% respectively. To sum up, multi-strategy parallel genetic algorithm based on machine learning K-means clustering is suitable for solving the workshop layout problem.

Table 7. Comparison of the results of three schemes

	Layout scheme	C_1/yuan	C_2	F	Cost savings /%	Optimizing comprehensive correlation degree/%
Original scheme	[5, 2, 3, 6, 4, 8, 7, 1]	536897.4	47.3	0.03498		
GA	[5, 2, 3, 8, 7, 4, 6, 1]	464922.2	47.9	0.01314	13.41	1.27
MLGA	[7, 5, 6, 2, 8, 1, 3, 4]	377217.5	50.3	-0.05572	29.74	6.34

5 Conclusion

This paper mainly discusses the workshop layout problem, studies the workshop equipment layout to ensure the non-logistics closeness of the workshop equipment while minimizing the cost, establishes a dual-objective optimization model of minimum handling cost and comprehensive correlation, and introduces a multi-strategy parallel genetic algorithm based on machine learning K-means (MLGA) clustering to expand the method to solve the workshop layout problem. Compared with the classical genetic algorithm, this algorithm is more stable and effective. In practice, From the example verification results, it can be seen that the optimization results of the proposed method are obviously better than those of the classical genetic algorithm, which verifies the effectiveness of the algorithm and lays a foundation for the future study of more complicated large-scale workshop layout problems.

References

1. Song, L. Xu, X. Sun, Y. (2010)A Hybrid Genetic Algorithm for Flexible Job Shop Scheduling . journal of management sciences in China, 13 (11): 49-

54.https://kns.cnki.net/kcms2/artcle/astract?v=w1Je9LIFm5BIpXf7GlmCi8NKYt7Gc1G_IOKiO3qzfxpfVBS_OLGMTeDz14R8DelR0oxE_vAK-TGuXqGc67Ub3SfSdzgufmqIt9l8OfVnZY5gO4TSat9UtyqYVoR5utyzSE7MTMbZU=&uniplatform=NZKPT&language=CHS

2. He, K. Shen, B. Peng, S. (2018) Workshop layout optimization of aerospace large-scale parts based on improved genetic algorithm. Modular machine tools and automatic processing technology , (01): 158-160. DOI: 10.13462/j.cnki.mmt.amt.2018.01.040.
3. Wan, J. Zhang, F. Gao, L. (2016) Study on Workshop Layout Planning Based on Genetic Simulated Annealing Algorithm [J]. Journal of Liaoning University of Technology (Natural Science Edition),,36 (01): 35-39.doi: 10.15916/j. issn 1674-3261.2016.01.
4. Dewi S ,Andriansyah ,Syahriza . (2019) Optimization of Flow Shop Scheduling Problem using Tabu Search Algorithm: A Case Study. IOP Conference Series: Materials Science and Engineering, 2019, 506 (1): 012064 (8pp).https://kns.cnki.net/kcms2/article/abstract?v=yqeyU9EK6jSHJs8iCiGZIUwYdzL4ijW1XVud2eUyGMmbw8yOZNlUgstTl51zZBSHkoL1Lyyuodl8TUGiXq5GSOMyPU5Lsxm55Oqyti3NxPG2MIVbbCMfLvJhtFP4dgFG4HgShWzIDLedhtEVUvJDTsz-zgsRmOgqm7pEMx8nb1c=&uniplatform=NZKPT&language=CHS
5. Tirkolaee B E ,Goli A ,Weber W G . (2020) Fuzzy Mathematical Programming and Self-Adaptive Artificial Fish Swarm Algorithm for Just-in-Time Energy-Aware Flow Shop Scheduling Problem with Outsourcing Option .IEEE Transactions on Fuzzy Systems, 2020, PP (99): 1-1.https://kns.cnki.net/kcms2/article/abstract?v=yqeyU9EK6jRGR_rEDhTh5_jsxKKxTCfdStgp0pwex28-LDLkvEek1O6hoDI5WCzD8-mVxComB6jqvH551NHxcfHQgxPSjWSrF6mRHYyEx9iz-TdGHlL5eOz6uDjcvZS6qeIPrnW1_FsaibQ0wBrrgmfTfRwD-dUDvRBY4bLtQ8_q0=&uniplatform=NZKPT&language=CHS
6. Ding, X. Yao, W. (2017) Multi-objective reconfigurable facility layout method based on particle swarm optimization . China Mechanical Engineering, 28 (07): 852-861.https://link.cnki.net/urlid/42.1294.th.20170405.1520.034
7. Lu ,Y. Li, X. Zhu, C. (2022)Multi-objective workshop layout optimization based on adaptive genetic simulated annealing algorithm. Manufacturing Technology and Machine Tool, (07): 173-179. DOI:10.19287/j.mtmt.1005-2402.2022.07.029.
8. Liu, X. Chen, X. Xi, S. (2022) Optimization method of multi-layer unit flow shop layout based on decomposition and coordination mechanism. Computer integrated manufacturingsystem,1-21.http://kns.cnki.net/kcms/detail/11.5946.TP.20220119.1620.010.html.
9. Zhang, S. Yu, D. Zheng, Y. (2022) Study on Workshop Layout Based on Multi-objective Migratory Bird Algorithm. Modern Manufacturing Engineering, (02): 16-23. DOI: 10.16731/j.cnki.1671-3133.2022.02.003.
10. Xu, Sha. (2022) Study on the warehouse layout and process optimization strategy of company X .China university of mining and technology. DOI: 10.27623/d.cnki.gzkyu.2022.000846.
11. Zhang, Y. Zhong, H. Zhang, C. (2021) Multi-strategy parallel genetic algorithm based on machine learning [J]. Computer Integrated Manufacturing System, 2021, 27 (10): 2921-2928. DOI:10.13196/j.cims.2021.10.016.