Aug. 2023 Transactions of Nanjing University of Aeronautics and Astronautics Vol. 40 No. 4

An Assembly Man-Hour Estimation Model Based on GA-SVM for Multi-specification and Small-Batch Production

XU Ji[1,2], *ZHANG Liping*[1], *LI Lu*[3], *XU Feng*[1*], *GUO Hun*[2*], *CHAO Haitao*[1],
ZUO Dunwen[1]

1. College of Mechanical & Electrical Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, P.R. China;
2. School of Aerospace & Mechanical Engineering, Changzhou Institute of Technology, Changzhou 213032, P.R. China;
3. Shanghai Spaceflight Precision Machinery Institute, Shanghai 201600, P.R. China

(Received 17 September 2022; revised 4 February 2023; accepted 21 July 2023)

Abstract: It is necessary to evaluate man-hour (MH) before receiving the order to guide the quotation and forecast the delivery date for a manufacturing contractor. As an important part of assembled MH, it has important practical significance. Aiming at the characteristics of multi-specification and small-batch production, an assembly MH estimation model based on support vector machine (SVM) is proposed. Apart from single component attributes, assembly process, and historical MH data, we also consider the average of shortest path length (ASPL), which quantifies the complexity of an assembly, as influencing factors of assembly MH. Furthermore, the auto calculating methods of these factors based on 3D models with Creo JLink are proposed. Through the comparison of several algorithms, SVM is chosen as the optimal algorithm for assembly MH modeling. Genetic algorithm (GA) is used to avoid the local solution and accelerate convergence when searching for the optimal parameters of SVM (c and g). Finally, the proposed GA-SVM model is trained and applied to predict the assembly MH of the bionic leg for the radar device. Experimental results show that GA-SVM has higher prediction accuracy than other methods in this paper and the whole predicting process only takes about 3 min.

Key words: assembly man-hour; multi-specification and small-batch; topological structure; genetic algorithm (GA); support vector machine (SVM)

CLC number: TH162 **Document code:** A **Article ID:** 1005-1120(2023)04-0500-11

0 Introduction

Delivery time, cost and quality are the primary challenges for companies today. The main goal of the company is to put the product on the market as soon as possible[1]. The manufacturer must estimate the time cost to ensure that the project is profitable[2]. For complex mechanical products, assembly man-hour(MH) account for almost 50% of the total MH, which is worth studying[3].

This study is based on the fact that the product process has not been formulated. As a part of the overall MH, the predicted assembly MH can be used as a reference for product quotation and delivery time, but cannot be used to arrange production.

Taylor[4] proposed the modular arrangement of the predetermined time standard, which has become an important part of the predetermined time standard. Park et al.[5] applied the depth camera to MH measurement and realized the automatic measurement of standard time. Mei et al.[6] proposed the idea of reasonably allocating labour in a single production to obtain the minimum working hours. Nitta et al.[7] has established a MH quota model, which can study errors according to the particularity of operators in the production process. Hanson et al.[8] studied the

---

*Corresponding author, E-mail addresses: xufeng@nuaa.edu.cn; guohun@nuaa.edu.cn.

How to cite this article: XU Ji, ZHANG Liping, LI Lu, et al. An assembly man-hour estimation model based on GA-SVM for multi-specification and small-batch production[J]. Transactions of Nanjing University of Aeronautics and Astronautics, 2023, 40(4): 500-510.

http://dx.doi.org/10.16356/j.1005-1120.2023.04.010

effect of matching tools on the MH of hybrid model assembly systems[8].

After the exploration for 50 years, the factors affecting MH have been gradually discovered and enriched. The focus of MH has turned to models and related algorithms. The methods for calculating the MH quota include mathematical model, empirical calculation, analysis method, statistical method, chart method, and simulation method. Tseng et al.[9] studied the case retrieval algorithm and the case modification technology, both of them belonged to the case-based reasoning technology and were applied to the calculation of MH in the processing process later. Chen et al.[10] used the expert evaluation method to conduct regression analysis on the similarity between the process and the corresponding working hours, and obtained the calculation model of working hours of the new similar process. Liu et al.[11] confirmed that the neural network algorithm was superior to the linear regression algorithm in the hull assembly MH prediction. Qu et al.[12] analyzed the factors affecting the assembly MH of ship sections and established the prediction model of assembly MH by using neural networks. Yu et al.[13] proposed a prediction model to estimate aircraft assembly MH by using support vector machine (SVM) optimized by particle swarm optimization.

In recent years, more and more experts model complex problems using algorithms. Song et al.[14] proposed a dynamic hybrid mechanism-based differential evolution algorithm to effectively schedule railway train delay. Deng et al.[15] proposed an enhanced MSIQDE algorithm for modeling complex problems. Zhao et al.[16] applied continuous wavelet transform and Gauss convolutional deep belief network to intelligent diagnosis. However, as far as we know, there are no studies using algorithms to predict assembly MH over the forecast period, as there is not known method to quantify the assembly complexity.

In terms of MH prediction algorithm, the linear regression algorithm is difficult to meet the highly nonlinear characteristics of MH. The neural network algorithm needs a large number of samples, which is difficult for enterprises marked as multiple varieties and small batches. SVM is used to small sample, nonlinear and multi value classification problems to a certain extent[17]. It overcomes the defects of “dimension disaster” and “excessive learning” in traditional machine learning methods such as neural networks[18].

Based on the historical data of similar products, this paper studies the influence of assembly relationship between parts of new products on assembly MH. The assembly attributes of products are extracted from the three-dimensional assembly model to solve the subjective uncertainty of workers. In addition, the shortest path method is used to quantify the assembly complexity and improve the assembly MH prediction model. Considering the problem of small historical samples and sparse data, the genetic algorithm (GA) is introduced to optimize the SVM, and the assembly MH prediction model is established.

1 Impacts of Assembly Man-Hours

There are many factors affecting assembly MH, including automation degree, component quality, assembly process, product structure, worker qualification, and product organization and management[19]. Therefore, the assembly MH quota belongs to a constrained mixed discrete non numerical optimization problem[20]. On the premise of the same level of assembly automation and worker proficiency, we mainly focus on the influence of part attributes, assembly process, and assembly topology.

1.1 Part attributes and assembly process

1.1.1 The way of component connecting

In addition to welded parts, components are a group of related parts connected together by a large number of connectors and fasteners, such as screws, washers, nuts, studs, etc[21]. The time of grasping, positioning, and other operations in the assembly process largely depends on the characteristics of the parts themselves.

The quality, surface area, volume and actual size of parts are important parameters that affect the assembly time. In consideration of the computational complexity and redundancy, part quality and bounding box size are selected as part attribute parameters that affect assembly MH. In the assembly process, the connection mode of components and the number of mating surfaces are mainly considered.

1.1.2 Identification of matching surface

The information of matching surfaces includes the number, the type, and the size of the matching surface. The larger the number of matching surfaces and the more area of matching surfaces are, the more time the assembly demands.

Chang et al.[22] presented a way to filter out the unrelated parts by determining whether two parts touched or intersected. Afterwards, assembly surfaces of parts that may have assembly relationships are matched. This paper proposes an assembly information extraction method with higher accuracy than the above method. The constraint information from 3D model is used to search for the couple of parts that may have an assembly relationship. Afterwards, surfaces of the couple are matched to judge whether the assembly relationship exists in the couple.

1.2 Assembly topology structure

We take the advantage of the complexity of assembly topology to quantify the difficulty of mechanical product assembly. According to the research of Mathieson et al.[23] and the lessons learned from graph theory, the assembled product can be expressed in the form of bipartite graph. Afterwards, the surfaces of the couple are matched to judge whether the assembly relationship is existed of the couple, and the process is shown in Fig.1.

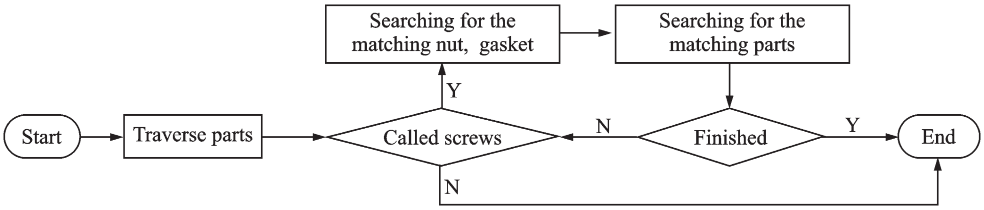


Fig.1 Identification of screw connecting

Fig.2 is a three-dimensional model of piston-connecting rod mechanism (PCRM) of motor, and Fig.3 is its form in bipartite graph.

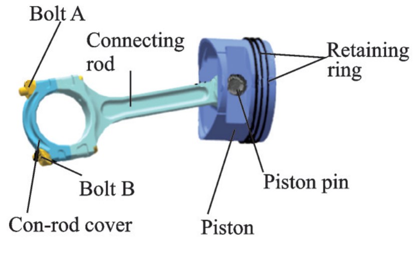


Fig.2 3D model of PCRM

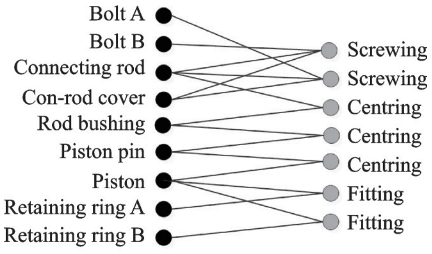


Fig.3 Bipartite graph of PCRM

Define a relational design structure matrix (rDSM)

$$\boldsymbol{D}_{nn}=\begin{bmatrix} d_{11} & \cdots & d_{1n} \\ \vdots & d_{ij} & \vdots \\ d_{n1} & \cdots & d_{nn} \end{bmatrix}$$

to mathematically express a product topology. Here n presents the total number of parts, and d_{ij} means whether part i and part j have a connection. If there are connectors between them, $d_{ij}=1$, otherwise, $d_{ij}=0$. The rDSM of above PCRM is

$$\boldsymbol{D}_{nn}=\begin{bmatrix} - & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & - & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & - & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & - & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & - & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & - & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & - & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & - & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & - \end{bmatrix}$$

The shortest path length refers to the minimum number of connections that an element passes to another given element in the system. It reflects the tightness of connections between components in the system. For example, between the connecting rod and the piston, the alignment of the connecting rod bushing, the alignment of the bushing and the piston pin, and the alignment of the piston pin and the piston must be determined. The shortest path length between the connecting rod and the piston is three.

The shortest path length between assembly parts can be calculated by Dijkstra algorithm[24]. Supposing a given bipartite graph $G = (V, E)$, where V represents the assembly element and E the connection relationship, we make $V = X + \overline{X}$ and $P_i \in X$, $i = 1, 2, \cdots, n$.

The basic idea of the algorithm is to first establish a set X to store points which have already calculated the shortest path and the set $\overline{X}$ stores the remaining points. The specific practices are as follows:

Step 1 Put the source point p_1 into X, and $P_i \in X$ represents the remaining $n-1$ points.

Step 2 Calculate the shortest path length m_{ij} between p_i (belongs to X) and p_j (belongs to $\overline{X}$), then put p_j into $\overline{X}$ and m_{ij} into all-pairs shortest path matrix (ASPM).

Step 3 Make $j+1$ and repeat Step 2 until all the points in V are put into set X.

Step 4 Make $i+1$ and make sure $i+j = n$, repeating Step 2 and Step 3 until $i=n$.

The ASPM of piston-connecting rod mechanism is

$$D^*_{nn} = \begin{bmatrix} 0 & 2 & 1 & 1 & 2 & 3 & 4 & 5 & 5 \\ 2 & 0 & 1 & 1 & 2 & 3 & 4 & 5 & 5 \\ 1 & 1 & 0 & 1 & 1 & 2 & 3 & 4 & 4 \\ 1 & 1 & 1 & 0 & 2 & 3 & 4 & 5 & 5 \\ 2 & 2 & 1 & 2 & 0 & 1 & 2 & 3 & 3 \\ 3 & 3 & 2 & 3 & 1 & 0 & 1 & 0 & 0 \\ 4 & 4 & 3 & 4 & 2 & 1 & 0 & 1 & 1 \\ 5 & 5 & 4 & 5 & 3 & 2 & 1 & 0 & 0 \\ 5 & 5 & 4 & 5 & 3 & 2 & 1 & 2 & 0 \end{bmatrix}$$

The average of the shortest path length was used to quantify the complexity of a product, shown as

$$D = \frac{\sum_{i}^{N_E} \sum_{j}^{N_E} D^*_{ij}}{N_E(N_E - 1)} \tag{1}$$

where N_E means the number of parts of a product and D the average value of the shortest path of PCRM.

The average value of the shortest path length reflects the degree of tightness between parts. Under the premise of avoiding the complexity of the part itself, the higher the value is, the tighter the relationship is between parts and the more difficult the assembly is. On condition that the number of parts is the same, as a result, the series structure has the largest average shortest assembly path value and the parallel structure has the smallest, which means that D has a certain range as follows

$$1 \leqslant D \leqslant \frac{\sum_{i=1}^{n-1} \sum_{j=1}^{i} j}{n(n-1)} \tag{2}$$

The average value of the shortest path length of Fig.2 is 2.67, that of Fig.4(a) is 1, and that of Fig.4(b) is 3.33.

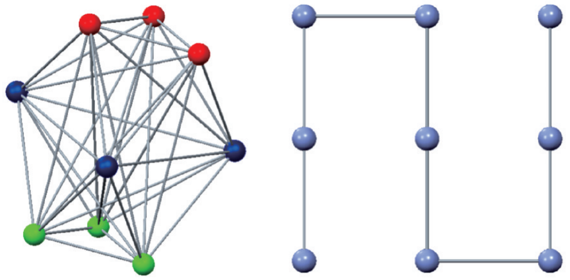

(a) Parallel structure
(b) Series structure

Fig.4 Topology diagram of assembly

2 Method for Predicting Assembly Man-Hour

This paper takes a bionic-leg structure from a specific machine as an example. The MH estimation process is shown in Fig.5.

2.1 Sample selection

Considering that the quality of the part is always proportional to the size of the part in the assembly of the bionic-leg structure, only the quality of the part is considered when considering the influence of the part characteristics on the assembly

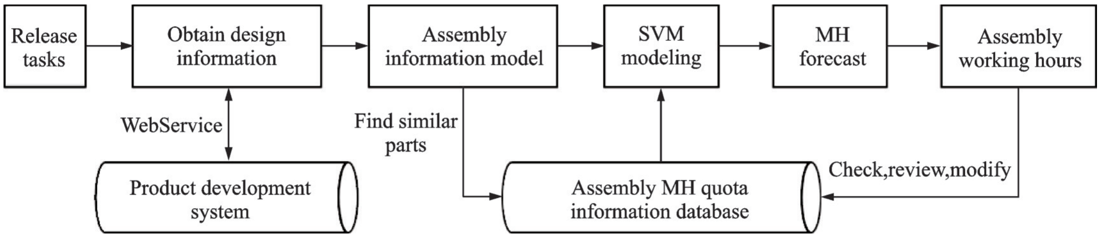


Fig.5 Assembly MH predication process of bionic-leg structure

time. According to the manufacturer's standard, parts with a mass of less than 1 kg are defined as small parts, and those with a mass of more than 25 kg are defined as large parts.

The bionic-leg structure is generally similar, usually using the same or similar assembly process. This paper selects 16 groups of historical data of bionic-leg assembly, and extracts the assembly information of new products in the three-dimensional model. The samples are shown in Table 1.

Table 1 Sample set

No.	P_1	P_2	P_3	P_4	A	L	C_1	C_2	M	T/h^{-1}
1	154	63	10	227	266	6.33	13	11	15	43
2	137	85	12	234	379	6.44	22	24	39	68
3	186	86	21	293	328	7.27	17	7	19	52
4	145	64	25	234	342	3.35	26	25	14	61
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
13	181	95	20	296	346	9.23	32	12	24	48
14	156	73	12	241	434	3.95	20	41	26	53
15	128	63	24	215	376	7.34	12	19	21	35
16	142	51	13	206	223	8.32	11	22	27	42

The quantities of small parts, medium parts, large parts and total parts are expressed as P_1, P_2, P_3, and P_4, respectively. Mass properties are defined in the Creo model by designer. So that P_1, P_2, P_3, and P_4 can be auto extracted by the development of the Creo Jlink application. A represents the number of assembly relationships and is auto calculated by extracting assembly info defined in the 3D model, and L the average length of the shortest path. C_1, C_2, and M represent the number of screws, pins, and adjacent relationships, which can also auto calculated by a Creo Jlink apploication. And T represents the assembly MH. The relationship between these variables can be described by the following mathematical expression as $T=[P_1, P_2, P_3, P_4, A, L, C_1, C_2, M]$.

Assuming that each influencing factor performs a one-dimensional linear fitting to the assembly MH, and making $y=ax+b$, the correlation coefficient is as shown in Table 2.

As can be seen from Table 2, the simple linear fitting cannot meet the prediction modeling of assembly MH as a result of the maximum R^2 being 0.4. The average length of the shortest path has the greatest influence on the univariate linear fitting coefficient of assembly MH because of the maximum absolute value a. The bigger the absolute value a is, the smaller the assembly MH is.

Table 2 Unitary linearity regress

Element	P_1	P_2	P_3	P_4	A	L	C_1	C_2	M
a	0.15	0.38	0.40	0.15	0.03	−3.85	0.81	0.25	0.51
b	32.17	25.9	48.88	16.25	45.52	79.63	37.88	49.24	43.61
R^2	0.14	0.21	0.03	0.24	0.02	0.40	0.35	0.05	0.12

2.2 Principle of SVM

MH quota is a typical highly nonlinear problem. In order to deal with the nonlinear problem, kernel function is introduced and the input data is mapped into high-dimensional linear feature space. Treat the P_1, P_2, P_3, P_4, A, L, C_1, C_1, M as input and T as output in SVM algorithm.

We describe the mathematical model $T=[P_1, P_2, P_3, P_4, A, L, C_1, C_2, M]$ as $y=[x_1, x_2, \cdots, x_n]$, where $n=9$.

Since there is no simple linear relationship between x_i and y, the training data set is mapped to a high-dimensional space utilizing $\varphi(x_i)$ and the decision function turns to the following form

$$y = \boldsymbol{\omega}^{\mathrm{T}} \cdot \varphi(x_i) + b \tag{3}$$

where $\boldsymbol{\omega}$ indicates the optimal hyperplane which classifies the factors affecting the assembly MH into different group and b the error offset. Mapping the input data into a high-dimensional linear space inevitably leads to dimensional disasters, and the computational difficulty increases dramatically.

The solution of Eq.(3) is similar to the solution of a linear regression problem, which is equivalent to solving

$$\begin{cases} \text{minimize} \quad \dfrac{1}{2}\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\omega} \\ \text{subject to} \quad y_i - (\boldsymbol{\omega} \cdot x_i + b) \leqslant \varepsilon \end{cases} \tag{4}$$

where ε represents the allowable error of the fitting of MH, and when the constraint cannot be satisfied, the slack variable ξ_i, ξ_i^* are introduced.

$$\begin{cases} \text{minimize} \quad \dfrac{1}{2}\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\omega} + c\sum_{i=1}^{l}(\xi_i + \xi_i^*) \\ \text{subject to} \begin{cases} y_i - \boldsymbol{\omega} \cdot \varphi(x_i) - b \leqslant \varepsilon + \xi_i \\ \boldsymbol{\omega} \cdot \varphi(x_i) + b - y_i \leqslant \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geqslant 0 \end{cases} \end{cases} \tag{5}$$

where c is the penalty coefficient indicating the degree of punishment of the algorithm for the wrong classified sample. Using the Lagrangian multiplication and the KKT condition, Eq.(5) is further deduced as

$$f(x) = \sum_{x_i \in \mathrm{SV}} (\alpha_i - \alpha_i^*) K(x_i, x_j) + b \tag{6}$$

where

$$b = \frac{1}{N_{\mathrm{NSV}}}\left\{ \sum_{0<\alpha_i<C}\left[y_i - \sum_{x_i \in \mathrm{SV}} (\alpha_j - \alpha_j^*) \cdot K(x_i, x_j) - \varepsilon \right] + \sum_{0<\alpha_i^*<C}\left[y_i - \sum_{x_i \in \mathrm{SV}} (\alpha_j - \alpha_j^*) K(x_i, x_j) + \varepsilon \right] \right\} \tag{7}$$

where $K(x_i, x_j)$ is the kernel function, introduced at reducing the amount of calculation directly and avoiding constructing a nonlinear mapping.

The common kernel functions used in SVM are polynomial functions, radial basis functions, sigmoid functions, and so on. Gaussian functions are a common form of radial basis functions, in the form as

$$K(x_i, x_j) = \exp((-x - y^2)/2\sigma^2) \tag{8}$$

The Gaussian function is one of the most widely used and most excellent kernel functions because of its wide converge domain and versatility. As described in Eq.(8), if σ is bigger enough, the weights of high-dimensional features will decay quickly, as a result they are actually equivalent to a low-dimensional subspace. By contrast, if σ is small enough, data can be mapped to be linearly separable. Based on these characteristics of Gaussian functions, it is chosen as the kernel function of SVM.

The model that is trained by SVM in this paper belongs to a multi-classification problem, in which different classifiers may have distinctive scales. For the purpose of solving this deficiency, the hyperplane with the largest interval and the corresponding b is calculated by Eq.(3). The assembly MH belongs to continuous variables, and classifiers are not suitable to predict it. In this paper, Eq.(3) is used to estimate assembly MH. The assembly MH studied is small-volume production enterprises, the size of the sample is small, and the imbalance of training samples will inevitably cause errors of the classifier to some extent.

2.3 Principle of GA-SVM

Searching for the appropriate c and g (g equals the kernel function parameter σ) is the key of SVM optimization, further to obtain b. At present, the commonly used grid crossover algorithm and empirical algorithm usually let c and g take discrete values in a certain range, and adopt the parameters with the best classification accuracy of the test set. There are certain deficiencies in efficiency and accuracy of both the grid crossover algorithm and the empirical optimization algorithm. GA is a random search algorithm which draws lessons from the natural selection and natural genetic mechanism of the biological world. It has the intrinsic characteristics of group search and heuristic random search, and it rarely falls into local optimum. The inherent parallel computing ability of GA can quickly find out the optimal values of c and g.

The input parameters should be normalized to eliminate the dimension and speed up the convergence in prior of training. Specific practices are as follows

$$y=\frac{(y_{\max}-y_{\min})(x-x_{\min})}{x_{\max}-x_{\min}}+y_{\min} \tag{9}$$

where $x_{\min}$ and $x_{\max}$ represent the minimum and maximum values of the input one-dimensional array. $y_{\max}$ and $y_{\min}$ represent the upper and lower limits of the normalized range of output MH, respectively. $y_{\max}=1$ and $y_{\min}=0$ are usually stipulated. The accuracy and adaptability of the algorithm are quantified by mean squared error (MSE) and R^2, respectively.

The normalized data are used as the input of LIBSVM, and the accuracy of cross validation algorithm is used as the fitness value. And c, g are treated as the parameters to be optimized. The flow chart is shown in Fig.6.

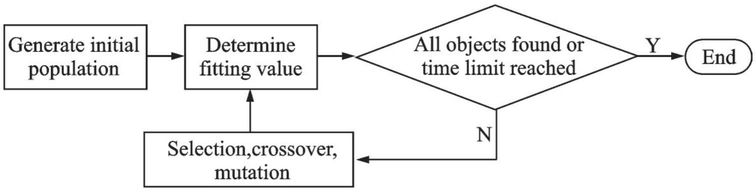


Fig.6 Flow of GA selecting SVM parameters

3 Example Verification

3.1 Results of examples

The first 13 groups of data in Table 1 are used as training sets, and the latter 3 groups are used to test the model precision.

In this case, the upper limit of evolutionary times of GA is 500, the number of population is 20, $c\in[0, 100]$, $g\in[0, 500]$, and the termination condition is that the number of iterations reaches half of the upper limit and the root-mean-square error is lower than a specified value.

Due to the randomness of the GA result is random, the performance of the algorithm cannot be fully illustrated by a single experiment. Repeated experiments are carried out under the same condition, whose results are shown in Table 3. In the case of similar MSE, c and g in the group with the smallest penalty coefficient are selected as the best parameters, and the mean value of the third group of data is taken. In the same accuracy rate, the smaller the penalty factor is, the stronger the prediction ability of the algorithm is.

As can be seen from Fig.7, the fitness is close to the present value and the algorithm has good convergence.

Table 3 Parameters selected from GA-SVM

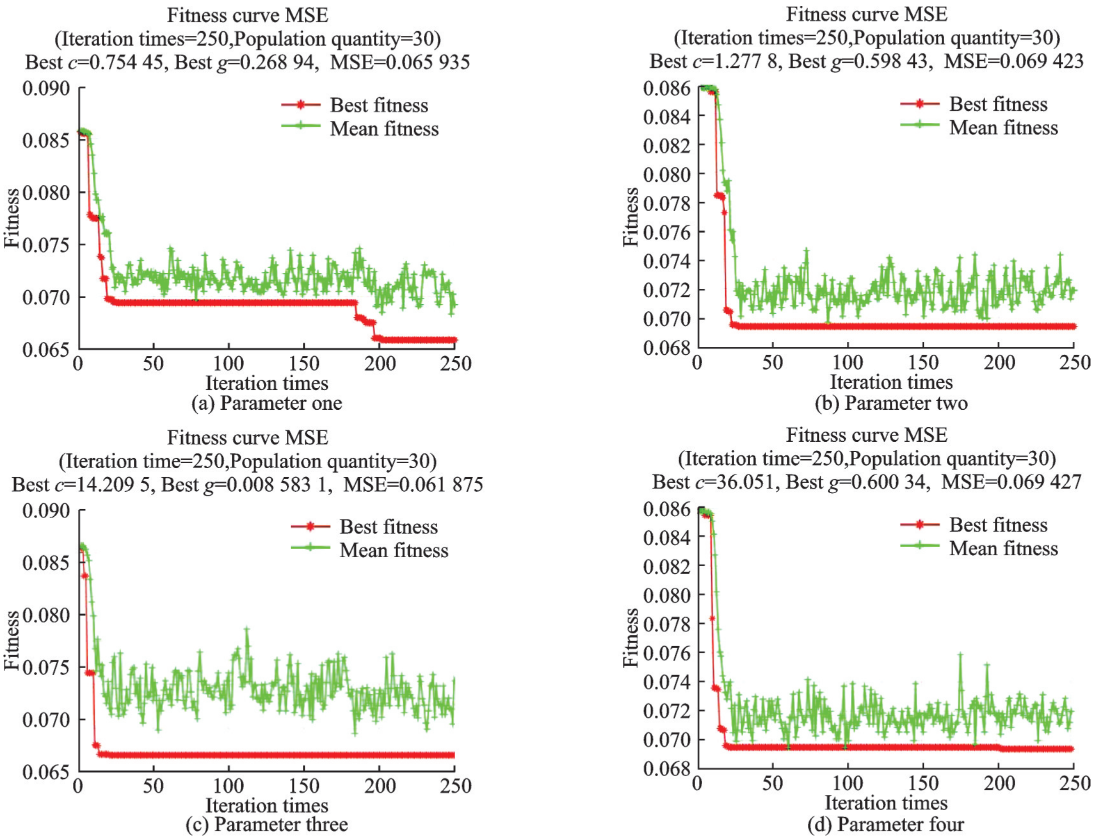
No.	Penalty coefficient c	Kernel parameter g	MSE	Squared correlation coefficient
1	0.75	0.27	0.066	0.916
2	1.13	0.16	0.064	0.908
3	1.28	0.60	0.069	0.999
4	1.48	0.60	0.069	0.999
5	1.79	0.61	0.069	0.999
6	3.85	0.61	0.069	0.999
7	5.20	0.60	0.069	0.999
8	7.52	0.61	0.069	0.999
9	21.92	0.61	0.069	0.999
10	36.05	0.60	0.069	0.999

In this paper, the mean impact value (MIV) is introduced into SVM based on GA. The main function of MIV is to rank the relativity between factors and outcomes.

We can conclude from Table 4 that the average shortest path has the greatest impact on the bionic-leg assembly MH, and the value of the average shortest path is inversely proportional to the assembly time.

3.2 Result analysis

Utilizing the established model and the third group data from Table 3 which means $c=1.28$, $g=0.60$, the assembly MH is predicted and the results are shown in Fig.8.



(a) Parameter one (b) Parameter two (c) Parameter three (d) Parameter four

Fig.7 Fitness curves of parameter selection

Table 4 Analysis with MIV

Parameter	P_1	P_2	P_3	P_4	A	L	C_1	C_2	M
MIV	0.000 4	0.019 2	0.002 5	0.006	0.001	0.042 6	0.008 6	0.004 8	0.031 1

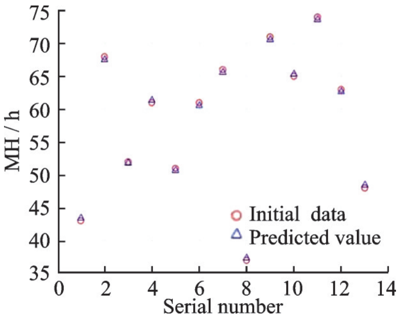


Fig.8 Predication result of algorithm

The MSE of the predicted bionic-leg assembly MH with GA-SVM is 0.036 and its square correlation coefficient is 0.999. The experiment proves that the algorithm has good accuracy and adaptability in predicting assembly MH with small samples.

In this paper, the grid optimization algorithm and GA are both used to predict assembly MH as well as BP neural network(BP-NN) algorithm. The experimental results verify the superiority of GA-SVM.

As shown in Table 5, the GA-SVM approach works best with the root mean square error of 3.669. The maximum error of prediction is 9.78%, and the minimum error is −3.66%. This meets the requirement of enterprise that the error of assembly MH is less than 30%.

4 Conclusions

For multi-specification and small-batch production, it is the key point to improve the core-competitiveness by reasonably estimating the cost and delivery time of products before subscribing an order. The assembly topology complexity is taken into account innovatively when estimating assembly

Table 5 Comparison of prediction results

Real value/h^{-1}	GA-SVM c=1.28, g=0.60		Grid research of SVM c=1.74, g=0.11		PSO-SVM c=12.02, g=0.01		BP-NN	
	Predicted value/h	Fractional error/%	Predicted value/h	Fractional error/%	Predicted value/h	Fractional error/%	Predicted value/h	Fractional error/%
64	61.648 0	−3.66	67.108 5	4.86	68.725 6	7.38	60.67	−5.20
45	49.398 8	9.78	48.033 5	6.74	49.164 4	9.25	52.04	15.64
53	50.430 1	−4.85	41.917 3	−21.73	45.398 4	−14.34	55.964 57	5.59
Root mean square error	—	3.669	—	4.86	—	7.38	—	5.2

MH which perfects the prediction model. All the influencing factors of assembly MH are auto extracted by the development of a Creo JLink application, and the whole predicting process takes only about 3 min. According to the scarcity and discretization of historical MH data, an assembly MH forecasting model is established by using GA-SVM. Verified by an example of bionic-leg structure, the root mean square error of GA-SVM is smaller than that of others referred in this paper. The forecasting error of GA-SVM is within 10%, which meets the enterprise requirement of 30% relaxation of general MH. The experiment also shows that the square correlation coefficient of MH prediction is 0.999, meaning the method has great adaptability.

References

[1] ELTAIEF A, LOUHICHI B, REMY S. Associations management and change propagation in the CAD assembly[J]. Computers in Industry, 2018, 98: 134-144.

[2] ISHII N, TAKANO Y, MURAKI M. An order acceptance strategy under limited engineering man-hours for cost estimation in Engineering-Procurement-Construction projects[J]. International Journal of Project Management, 2014, 32(3): 519-528.

[3] ZHOU X X, XIN B, GONG Y D, et al. Analysis of effective working hours of automatic assembly equipment for aircraft assembly stations[J]. Transactions of Nanjing University of Aeronautics and Astronautics, 2020, 37(6): 903-913.

[4] TAYLOR D H. Global cases in logistics and supply chain management[M]. [S.l.]: Cengage Learning EMEA, 1997.

[5] PARK J W, KIM D Y. Standard time estimation of manual tasks via similarity measure of unequal scale time series[J]. IEEE Transactions on Human-Machine Systems, 2017, 48(3): 241-251.

[6] MEI Y, ZENG Z, FENG D, et al. A method for man hour optimization and workforce allocation problem with discrete and non-numerical constraints in large-scale one-of-a-kind production[J]. International Journal of Production Research, 2016, 54(3): 864-877.

[7] NITTA I, SHIBUYA T, HOMMA K. Method and apparatus for estimating man-hours: US20090055142 A1[P]. 2009-02-26.

[8] HANSON R, MEDBO L. Aspects influencing man-hour efficiency of kit preparation for mixed-model assembly[J]. Procedia CIRP, 2016, 44: 353-358.

[9] TSENG H E, CHANG C C, CHANG S H. Applying case-based reasoning for product configuration in mass customization environments[J]. Expert Systems with Applications, 2005, 29(4): 913-925.

[10] CHEN Youling, REN Xiaojie. Time-quota computing method based on similarity coefficient of process[J]. Computer Integrated Manufacturing System, 2014, 4: 866-872. (in Chinese)

[11] LIU B, JIANG Z. The man-hour estimation models & its comparison of interim products assembly for shipbuilding[J]. International Journal of Operations Research, 2005, 2(1): 19-14.

[12] QU Shipeng, JIANG Zuhua. Man-hour calculation methods of the block assembly for shipbuilding[J]. Journal of Harbin Engineering University, 2012, 33: 550-555. (in Chinese)

[13] YU T, CAI H. The prediction of the man-hour in aircraft assembly based on support vector machine particle swarm optimization[J]. Journal of Aero-space Technology and Management, 2015, 7: 19-30.

[14] SONG Y, CAI X, ZHOU X, et al. Dynamic hybrid

mechanism-based differential evolution algorithm and its application[J]. Expert Systems with Applications, 2023, 213: 118834.

[15] DENG W, XU J, GAO X Z, et al. An enhanced MSIQDE algorithm with novel multiple strategies for global optimization problems[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2020, 52(3): 1578-1587.

[16] ZHAO H, LIU J, CHEN H, et al. Intelligent diagnosis using continuous wavelet transform and Gauss convolutional deep belief network[J]. IEEE Transactions on Reliability, 2022. DOI: 10.1109/TR. 2022. 3180273.

[17] SUN Shengshan, XU Changkai, HE Yaqun. Prediction model of air material consumption based on RS-PSO-SVM[J]. Journal of Nanjing University of Aeronautics & Astronautics, 2021, 53(6): 881-887. (in Chinese)

[18] PENG X, WANG Y. A normal least squares support vector machine (NLS-SVM) and its learning algorithm[J]. Neurocomputing, 2009, 72(16/17/18): 3734-3741.

[19] LIU Bin, JIANG Zuhua. The intelligent man-hour estimate technique of assembly for shipbuilding[J]. Journal of Shanghai Jiao Tong University, 2005, 39(12): 1979-1983. (in Chinese)

[20] HADJ R B, BELHADJ I, TRIGUI M, et al. Assembly sequences plan generation using features simplification[J]. Advances in Engineering Software, 2018, 119: 1-11.

[21] MATHIESON J L, SUMMERS J D. Complexity metrics for directional node-link system representations: Theory and applications[C]//Proceedings of International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. [S.l.]: [s.n.], 2010, 44137: 13-24.

[22] CHANG C C, LIN C J. LIBSVM: A library for support vector machines[J]. ACM Transactions on Intelligent Systems and Technology (TIST), 2011, 2(3): 1-27.

[23] MATHIESON J L, WALLACE B A, SUMMERS J D. Assembly time modeling through conective complexity metrics[J]. International Journal of Computer Integrated Manufacturing, 2013, 26(10): 955-967.

[24] WANG Shuxi, LI Anyu. Multi-adjacent-vertexes and multi-shortest-paths problem of Dijkstra algorithm[J]. Computer Science, 2014, 41(6): 217-224. (in Chinese)

Acknowledgements This work was supported in part by the Special Fund of Jiangsu Province for the Transformation of Scientific and Technological Achievements (Nos.BA2018110, BA2021036), and the Natural Science Foundation of Jiangsu Province (No.BK20211061). The authors would like to acknowledge the following people for their assistance: Mr. YU Ke, Mr. ZHOU Guoping, Mr. HE Yafeng, all with Changzhou Wujin GUANGYU Embossing Roller Machinery Co., Ltd.

Authors Ms. **XU Ji** received the B.S. and M.S. degrees in mechanical engineering from Hohai University, in 2005 and 2012, respectively. From 2018, she is studying for a doctorate at Nanjing University of Aeronautics and Astronautics, mainly engaged in fault diagnosis and intelligent processing technology research.

Prof. **XU Feng** received the B.S. and Ph.D. degrees in mechanical engineering from Nanjing University of Aeronautics and Astronautics in 1999 and 2007, respectively. His research is focused on superhard coating and advanced manufacturing technology research.

Prof. **GUO Hun** received the B.S. and Ph.D. degrees in mechanical engineering from Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 1998 and 2006, respectively. His research is focused on intelligent machining technology and advanced manufacturing technology research.

Author contributions Ms. **XU Ji** designed the study, complied the models, conducted the analysis, interpreted the results and wrote the manuscript. Ms. **ZHANG Liping** conducted experiments, wrote the manuscript and searched literatures. Mr. **CHAO Haitao** and Ms. **LI Lu** contributed to data and model components for the experiment. Prof. **XU Feng** conducted supervision, technical route guidance and fund acquisition. Prof. **GUO Hun** contributed to overall planning and design of the project. Prof. **ZUO Dunwen** contributed to guidance and help. All authors commented on the manuscript draft and approved the submission.

Competing interests The authors declare no competing inrests.

(Production Editor: XU Chengting)

基于GA-SVM的多规格小批量生产的装配工时估算模型

徐 吉[1,2]，张丽萍[1]，李 露[3]，徐 锋[1]，郭 魂[2]，晁海涛[1]，左敦稳[1]

（1.南京航空航天大学机电学院，南京 210016，中国；2.常州工学院航空航天与机械工程学院，常州 213032，中国；3.上海航天精密机械研究所，上海 201600，中国）

摘要：对于制造承包商来说，在正式接收订单之前，为了指导报价和预测交货日期，有必要对工时（Man-hours，MH）进行评估。装配工时作为工时的重要组成部分，具有重要的实际研究意义。针对多规格、小批量生产的特点，提出了一种基于支持向量机（Support vector machine，SVM）的装配工时估算模型。除了单部件属性、装配过程和历史工时数据外，还考虑了可量化装配复杂性的最短路径长度平均值（Average of shortest path length，ASPL）作为装配MH的影响因素，并提出了基于Creo JLink三维模型的这些因素的自动计算方法。通过对几种算法的比较，选择SVM作为装配体MH建模的最优算法。将遗传算法（Genetic algorithm，GA）应用于SVM中，有利于在SVM中搜索最优参数c和g时避免了局部求解，加快了收敛速度。最后，对所提出的GA-SVM模型进行训练，并应用于雷达装置仿生腿的装配工时预测。实验结果表明，GA-SVM具有比本文其他方法更高的预测精度，整个预测过程仅需3 min左右。

关键词：装配工时；多规格小批量；拓扑结构；遗传算法；支持向量机