EAI Endorsed Transactions
on Scalable Information Systems
Research Article EAI.EU

Intuitionistic Fuzzy Set Similarity Degree Based on Modified Genetic Algorithm for Solving Heterogenous Multi-dimension Targeted Poverty Alleviation Data Scheduling Problem

Yang Shi[1], Qingwu Shi[1,*] and Xiaofeng Mou[1]

[1]College of Information and Electronic Technique, Jiamusi University, Jiamusi, 154007 China
Email: softwarediy@126.com;shiqingwu@126.com;352720214@qq.com

Abstract

Targeted poverty alleviation is a proposed concept in comparison with extensive poverty alleviation. It mainly aims at the poverty situation of different rural areas and farmers in China and adopts scientific and reasonable methods to carry out targeted assistance policies. It executes accurate management for the targeted poverty alleviation. This way for poverty alleviation is more precise. In the research of heterogenous multi-dimension targeted poverty alleviation data scheduling, the multi-dimension processing is very important. In this paper, we propose an intuitionistic fuzzy set similarity degree based on modified genetic algorithm for solving heterogenous multi-dimension targeted poverty alleviation data scheduling problem. In the proposed algorithm, the reference solution and Pareto solution are mapped to the reference solution intuitive fuzzy set and Pareto solution intuitive fuzzy set respectively. The intuitionistic fuzzy similarity between two sets is calculated to judge the quality of Pareto solution. The similarity value of intuitionistic fuzzy sets is used to guide the evolution of multi--dimension genetic algorithm. The results show that the proposed algorithm can effectively solve the problem of heterogenous multi-dimension targeted poverty alleviation data scheduling, especially, in large scale problems.

Keywords: heterogenous multi-dimension targeted poverty alleviation, intuitive fuzzy set, genetic algorithm, Pareto solution.

Received on 12 October 2021, accepted on 22 October 2021, published on 22 October 2021

doi: 10.4108/eai.22-10-2021.171597

*Corresponding author. Email: shiqingwu@126.com

1. Introduction

Rural poverty has always been an important factor hindering China rural economic development. It is an important task for China rural poverty alleviation and development work to help the poor people in rural areas get rid of poverty smoothly. But poverty control work is still very difficult to carry out, which has brought great pressure to the party and the government. This is because poverty alleviation data scheduling is not timely and data cannot be shared in real time, which affects the accuracy of poverty alleviation work.

Data scheduling problems cover a wide range including single machine scheduling problem, resource scheduling problem, poverty alleviation job scheduling problem and open data scheduling problem [1]. About 25% of the production and manufacturing systems can be simplified to data scheduling problem. Therefore, the multi-dimension targeted poverty alleviation data scheduling problem (MTPA) is one of the most widely studied production scheduling problems. For the single-dimension targeted poverty alleviation data scheduling problem, Wang [2] proposed an evolutionary search algorithm based on multiple agents. Zheng [3] used the improved basic quantum evolutionary algorithm to solve the problem. Lin [4] proposed a hybrid-discrete biogeography based optimization (HDBBO) algorithm to solve the problem. Liu [5] proposed a hybrid discrete artificial bee colony algorithm. Lu [6] proposed a hybrid multi-objective backtracking search algorithm to solve the multiple targets flow-shop scheduling problem. Li [7] proposed a multi-population and multi-objective genetic algorithm based on decomposition method. At present, for multi-dimension targeted poverty alleviation data scheduling, one or two optimization objectives are mostly considered. Considering more than three dimensions targeted poverty alleviation data scheduling problems is less because of the complexity of targeted poverty alleviation data scheduling problem [8,9]. On the other hand, with the increase of the target dimension, the proportion of the non-dominant solution increases rapidly when the commonly used multi-dimension targeted poverty alleviation data scheduling algorithm is adopted. It is difficult to distinguish the better solution with the dominance relation as the fitness value resulting in the high-dimensional multi-objective (target number). Multi-dimension targeted poverty alleviation data scheduling is extremely challenging, but with the increasingly fierce competition among enterprises, decision makers need to consider a variety of factors when selecting scheduling schemes, such as completion time, delay cost, idle time, cost, inventory, etc,. Therefore, it is particularly important to study the high-dimensional multi-objective data scheduling problem.

Genetic algorithms (GA) are derived from computer simulations of biological systems [10]. Compared with other algorithms, genetic algorithm has the unique characteristics of universality, adaptability, expansibility and implicit parallelism, and has great advantages in solving the approximate optimal solution of complex multi-dimension targeted poverty alleviation data scheduling problems. After years of research and experiments, it is proved that the genetic algorithm is feasible and effective in solving the heterogenous multi-dimension targeted poverty alleviation data scheduling problem.

Intuitionistic fuzzy set is a generalization of fuzzy set theory, which considers membership degree, non-membership degree and hesitation degree at the same time. Many scholars have studied the measurement analysis of fuzzy set and applied it to various fields. Nguyen [11] proposed a similarity/dissimilarity measurement method for intuitionistic fuzzy sets and applied it to pattern recognition. Ananthi [12] proposed a thresholding method based on interval intuitionistic fuzzy sets and applied it to image segmentation. Lee [13] put forward a kind of intuitionistic fuzzy genetic algorithm solving weapon target assignment problem. Wang kaijun combined intuitionistic fuzzy set and grey model in fuzzy mathematics to design a fault prediction method. At present, the application of intuitionistic fuzzy sets similar to the research of shop scheduling is few.

MTPA is a simplified model for many practical production scheduling problems. Therefore, this paper combines intuition fuzzy similarity method with modified genetic algorithm and proposes a SIFSMGA algorithm, which is applied to multi-dimension targeted poverty alleviation data scheduling problem. The similarity of intuitionistic fuzzy set is used to explore the effective information between Pareto solution and each target, effectively eliminating the influence of magnitude and dimension of target, solving the problem of information loss [14]. The main contributions are as follows. The similarity of intuitionistic fuzzy sets is used as the fitness of the evolutionary algorithm to guide the evolution of the algorithm. Pareto solution is mapped to intuitionistic fuzzy set by membership function, non-membership function and intuitionistic index. It calculates the similarity between Pareto solution intuitionistic fuzzy set and reference solution intuitionistic fuzzy set. The advantage of Pareto solution is judged by the similarity size. Simulation experiments are conducted on 6 CEC standard test sets and 10 multi-dimension targeted poverty alleviation data scheduling test cases. The superiority of SIFSMGA algorithm is verified, and it can effectively solve the multi-dimension targeted poverty alleviation data scheduling problem, especially in solving large-scale problems.

2. Mathematical model

The MTPA problem can be regarded as a job shop scheduling problem. Therefore, the description of MTPA problem is as follows:

n workpieces are processed on m machines. The processing path of each workpiece is the same, and the processing order of the workpiece on all machines is the same. The processing time of one workpiece on one machine is fixed. The same workpiece cannot be processed on different machines at the same time, and multiple workpieces cannot be processed on one machine at the same time. The goal is to study and determine the optimal processing sequence of all the workpiece so as to achieve the optimal set of performance indicators. Where, the set of n workpieces is $J = \{1,2,\cdots,n\}$, the set of m machines is $m = \{1,2,\cdots,m\}$.

After receiving the order from the retailer, the manufacturer needs to process n workpieces and transport them to the retailer after processing. Subjected to the retailer's delivery date and its own inventory capacity, the manufacturer must complete all work within a reasonable time. The completion of the work ahead of time or delay will have a negative impact on the economic benefits of the enterprise. The workpieces need to be stored before they can be transported. If the workpiece is completed in advance, there will be inventory costs; If the workpiece is completed after the delivery date, there will be delay cost. Inventory cost and delay cost are proportional to time. After the work is finished, it will be transported to the retailer in batches. The maximum volume of shipments is fixed and the same. When the number of finished workpieces reaches the maximum traffic, they will be transported. And the transporting time of each batch is the completion time of the last workpiece. If the last batch of workpieces is less than the maximum shipment, the last batch of workpieces will still be transported as a batch.

The problem can be described by the following notations. $x=\{x_1,x_2,\cdots,x_n\}^T$ is the vector of n decision variables, which denotes a feasible workpiece processing sequence. $f_k(x)$ is the function value of the k-th objective function. $F(x)=(f_1(x)f_2(x)\cdots f_k(x)\cdots f_M(x))^T$ is the target vector or Pareto front edge. $Z^*=(F^*)=(f_1^* f_2^*\cdots,f_k^*\cdots f_M^*)^T$ is the ideal target vector. The k-th component of the ideal target vector is the minimum value of the following problem: $f_k^*=\min f_k(x)$. Other parameters are as follows:

n : total number of workpieces.

m : total number of machines.

T_{ik} : the processing time of workpiece i on machine k. i=2,…,n; k=2,…,m;

$C_{im}(x)$: in a sequence of workpiece, the completion time of workpiece i on machine m.

D_i : the delivery time of the i-th workpiece.

L_i : the delivery time of the i-th workpiece.

β_i : the inventory cost per unit of time for the i-th workpiece.

σ_i : the delay cost per unit of time for the i-th workpiece.

$S(j)$: the actual traffic of i-th batch.

h : maximum traffic of single batch.

b : delivery times.

In enterprises, different departments have different expectations for shop floor scheduling decisions based on their own interests. For example, the production workshop wants to improve production efficiency, the sales department wants to be able to deliver goods on time, and the manufacturing department wants to reduce costs as much as possible. Therefore, the following meaningful and conflicting optimization goals are set in this paper. The purpose of shortening the maximum completion time is to improve the production efficiency of the enterprise. The purpose of reducing the maximum delay time is to improve customer satisfaction. Reducing the delay cost and inventory cost is to improve the economic efficiency of enterprises. The optimized objective function is as follows:

The completion time (i.e., the maximum completion time) of the last workpiece on the last machine is:

$$f_1(x)=\max\{C_{im}(x)\mid i\in 1,2,\cdots,n\} \tag{1}$$

Because of the delay in completion of each workpiece, the manufacturer will be charged (i.e., the maximum delay time) is:

$$f_2(x)=\max\{(0,(C_{im}(x)-D_i))\mid i\in 1,2,\cdots,n\} \tag{2}$$

The inventory cost (i.e., the total inventory cost) of each workpiece before transportation is:

$$f_3(x)=\sum_{i=1}^{n}(\beta_i\cdot\max\{0,(L_i-C_{im}(x))\}) \tag{3}$$

A certain delay cost (i.e., total delay cost) may be incurred for all the delayed workpieces, namely

$$f_4(x)=\sum_{i=1}^{n}(\sigma_i\cdot\max\{0,(C_{im}(x)-D_i)\}) \tag{4}$$

Completion time $C_{ik}(x)$ and objective function are subject to the following constraints:

$$C_{11}(x)\ge T_{11} \tag{5}$$

$$C_{1k}(x)\ge C_{1(k-1)}(x)+T_{11}, k=2,\cdots,m \tag{6}$$

$$C_{i1}(x)\ge C_{1(i-1)}(x)+T_{i1}, k=2,\cdots,n \tag{7}$$

$$\begin{aligned}&C_{ik}(x)\ge\max\{C_{i(k-1)}(x),C_{k(i-1)}(x)\}+T_{ik},\\&k=2,\cdots,m;i=2,\cdots,n\end{aligned} \tag{8}$$

$$S(j)\le h, j=1,2,\cdots,b \tag{9}$$

Formula (9) describes the transport capacity of all transports. The mathematical model of the multi-objective replacement flow shop scheduling problem is as follows:

$$F(x)=\min\{f_1(x),f_2(x),f_3(x),f_4(x)\} \tag{10}$$

3. An improved genetic algorithm based on similarity of intuitionistic fuzzy sets

We adopt the similarity of intuitionistic fuzzy sets as the fitness of evolutionary algorithms to guide the evolutionary algorithm. Through the membership functions, the non-membership functions and intuitionistic index, it maps the Pareto solution and reference solution to form the Pareto solution intuitionistic fuzzy sets and reference solution intuitionistic fuzzy sets. Calculate the similarity between two intuitionistic fuzzy sets, in a similar size to judge the merits of the Pareto solutions. The similarity between two intuitionistic fuzzy sets is calculated to judge the quality of Pareto solution.

3.1. Definition of intuitionistic fuzzy sets

Intuitionistic Fuzzy Set (IFS) is an extension of the fuzzy set theory to deal with uncertain incomplete information. It uses the membership degree, non-membership degree and hesitation (intuitionistic index) to describe information to evaluate the relationship between comparison fuzzy sets. The intuitionistic fuzzy set is defined as: If X is a given finite field of element x, then an intuitionistic fuzzy set A on X is,

$$A = \{x, \mu_A(x), \gamma_A(x) \mid x \in X\} \tag{11}$$

Where, the membership function of intuitionistic fuzzy set A is $\mu_A(x): X \to [0,1]$. The non-membership function of intuitionistic fuzzy set A is $\gamma_A(x): X \to [0,1]$. For all the $x \in X$ in A, $\mu_A(x) + \gamma_A(x) \le 1$.

In order to measure the hesitation degree of x for A, in each intuitionistic fuzzy subset of x, we put forward intuitionistic index of x.

$$\pi_A(x) = 1 - \mu_A(x) - \gamma_A(x) \tag{12}$$

If $\pi_A(x) = 0$, $\forall x \in X$, then $\mu_A(x) + \gamma_A(x) = 1$. The intuitionistic fuzzy set A becomes the ordinary fuzzy set.

If the totality of intuitionistic fuzzy sets defined on X is represented by IFS(X), then an intuitive fuzzy set A∈IFS(X), its membership degree $\mu_A(x)$, non-membership degree $\gamma_A(x)$ and intuitive index $\pi_A(x)$ respectively represent the degree of supporting, opposing and neutral evidence of the intuitionistic fuzzy set A.

3.2. Set up intuitionistic fuzzy sets

When applying the of intuitionistic fuzzy sets similarity to multi-objective optimization, it is necessary to establish a relation between the multi-objective solution and the intuitionistic fuzzy sets. For this reason, this paper proposes a two-step mapping of multi-objective solution intuitionistic fuzzy set. Each sub-objective is optimized with q order single objective, and the maximum value of q order single objective optimization is selected as the upper bound of the sub-objective theory domain, and the λ times average value of the optimal value of the q order single objective optimization is selected as the lower bound of the sub-objective theory domain.

Step 1.

Mapping each sub-objective solution to an intuitionistic fuzzy subset. Define the mapping H_1:

$$f_i^j(x) \xrightarrow{H_1} \{x, \mu_{f_i}(j), \gamma_{f_i}(j)\} \tag{13}$$

Where $x = [x_1, x_2, \cdots, x_n]$ represents a feasible workpiece processing sequence. $i = 1,2,\cdots,M$. M is the number of objective functions. $j = 1,2,\cdots,N$. N is the number of solutions. $f_i^j(x)$ is the i-th sub-objective function value of the j-th solution. $\mu_{f_i}(j)$ is the membership value of the i-th sub-objective function of the j-th solution, indicating the similarity degree. $\gamma_{f_i}(j)$ is the non-membership value of the i-th sub-objective function of the j-th solution, indicating the degree of dissimilarity.

In here, the membership function of i-th sub-objective function $f_i^j(x)$ of the j-th solution is as follows:

$$\mu_{f_i}(j) = \begin{cases} 1, & f_i^j(x) \le f_i^{\min} \\ \dfrac{f_i^{\max} - f_i^j(x)}{f_i^{\max} - f_i^{\min}}, & f_i^{\min} < f_i^j(x) < f_i^{\max} \\ 0, & f_i^j(x) > f_i^{\max} \end{cases} \tag{14}$$

In the formula, $f_i^{\max}$ and $f_i^{\min}$ are the upper and lower bounds of sub-target i on the field U, namely, the range of each sub-target value.

Genetic algorithm is used to optimize sub-objective i, and the average value of objective function $f_i^0(x)$ after q-order optimization is obtained. We get the lower bound: $f_i^{\min} = \lambda f_i^0(x)$, $\lambda \in (0,1)$. We take the upper bound $f_i^{\max}$ as the maximum value of the first generation sub-target in all q optimizations.

The non-membership function of the i-th sub-objective function $f_i(x)$ of the j-th solution is defined as follows:

$$\gamma_{f_i}(j) = \begin{cases} 0, & f_i^j(x) \le f_i^{\min} \\ \dfrac{f_i^j(x) - f_i^{\min}}{\alpha_i f_i^{\max} - f_i^{\min}}, & f_i^{\min} < f_i^j(x) < f_i^{\max} \\ 1, & f_i^j(x) > f_i^{\max} \end{cases} \tag{15}$$

Where $\alpha_i \in [1,1.3]$ is a parameter.

Step 2.

The set of each sub-target intuitionistic fuzzy subset is mapped to a multi-objective intuitionistic fuzzy set. According to equations (14) and (15), the intuitionistic fuzzy subset of the i-th sub-target of the j-th solution is calculated. Define the mapping H2 as:

$$\begin{aligned} &Y(f_1^j(x), \cdots, f_i^j(x), \cdots, f_M^j(x)) \xrightarrow{H_2} \\ &IFS(A_{f_1}^j, \cdots, A_{f_i}^j, \cdots, A_{f_M}^j) \end{aligned} \tag{16}$$

Where $A_{f_i}^j$ is the intuitionistic fuzzy subset of the i-th objective of the j-th multi-objective solution.

To sum up, in order to establish the relationship between multi-objective solutions and intuitionistic fuzzy sets, H1 and H2 map are needed. By mapping H1, the j-th solution is calculated for each sub-target in intuitionistic fuzzy subset.

By mapping H2, M intuitionistic fuzzy subsets obtained by mapping H1 are mapped to intuitionistic fuzzy sets IFS. The mapping process is as follows:

$$Y(f_1^j(x),\cdots,f_i^j(x),\cdots,f_M^j(x)) \rightarrow \begin{cases} f_1^i(x) \xrightarrow{H_1} \{x,\mu_{f_1}(j),\gamma_{f_1}(j)\}, \\ \cdots \\ f_i^i(x) \xrightarrow{H_1} \{x,\mu_{f_i}(j),\gamma_{f_i}(j)\}, \quad \xrightarrow{H_2} \\ \cdots \\ f_M^i(x) \xrightarrow{H_1} \{x,\mu_{f_M}(j),\gamma_{f_M}(j)\} \end{cases}$$

$$IFS(x,\mu_{f_1}(j),\gamma_{f_1}(j),\cdots,\mu_{f_i}(j),\gamma_{f_i}(j),\cdots,\mu_{f_M}(j),\gamma_{f_M}(j)) = IFS(A_{f_1}^j,\cdots,A_{f_i}^j,\cdots,A_{f_M}^j)$$

(17)

In the multi-objective optimization problem, there is often a contradiction between each target. Generally, there is no ideal solution to make each target reach its own optimal value. Therefore, in the target space, this paper conducts q single-target optimizations for each target, and takes the average value of q single-target optimal values as the point, which will be as the reference solution of the multi-objective problem. In order to solve the similarity between intuitionistic fuzzy sets, the reference solution and Pareto solution should be mapped into intuitionistic fuzzy sets first.

1) Reference solution mapping

Genetic algorithm is used to optimize sub-objective i, and the average value of q target values after q optimizations is taken to form the multi-objective reference solution function value $Y_0 = (f_1^0(x), f_2^0(x),\cdots,f_M^0(x))$. Using the above two-step mapping, Y_0 is mapped to the reference solution intuitionistic fuzzy set $IFS(0) = (A_{f_1}^0,\cdots,A_{f_i}^0,\cdots,A_{f_M}^0)$. Where $f_i^0(x)$ is the average function value after of the i-th sub-objective optimization. $A_{f_i}^0$ is the reference solution intuitionistic fuzzy set after the i-th sub-target mapping, $A_{f_i}^0 = \{(x,\mu_{f_i}(0),\gamma_{f_i}(0))\}$.

2) Pareto solution mapping

After each iteration of multi-objective genetic algorithm, the current Pareto optimal frontier $Y_j = (f_1^j(x), f_2^j(x),\cdots,f_M^j(x))$. Through the above two-step mapping, Y_j is mapped to Pareto to solve the intuitionistic fuzzy set $IFS(j) = (A_{f_1}^j,\cdots,A_{f_i}^j,\cdots,A_{f_M}^j)$. Where $f_i^j(x)$ is the i-th sub-object function value of the j-th Pareto solution. $A_{f_i}^j$ is Pareto intuitive fuzzy set after the i-th sub-object mapping of the j-th Pareto solution, $A_{f_i}^j = \{(x,\mu_{f_i}(j),\gamma_{f_i}(j))\}$.

3.3. Similarity measure between intuitionistic fuzzy sets

In order to calculate the similarity between Pareto solution and reference solution, it is necessary to measure the similarity of intuitionistic fuzzy set after mapping. Reference [15] put forward an intuitionistic fuzzy sets measurement method according to the geometric interpretation of IFS. This paper proposes a distance measurement formula between multi-objective intuitionistic fuzzy sets as (18):

$$q(IFS(j),IFS(0)) = \sqrt{\frac{1}{2M}\sum_{i=1}^{M}[(\mu_{f_i}(j)-\mu_{f_i}(0))^2 + (\gamma_{f_i}(j)-\gamma_{f_i}(0))^2 + \rho\cdot(\pi_{f_i}(j)-\pi_{f_i}(0))^2]} \quad (18)$$

Where $j = 1,\cdots,N$. N is the number of solutions. The ρ value considers the indecision state, which contains similarity and dissimilarity. In the absence of any other prior information, it is reasonable to consider that similarity and dissimilarity are equal to each other, so ρ=0.5. The similarity between intuitionistic fuzzy sets with multi-objective is calculated according to the distance measurement.

$$S(IFS(j),IFS(0)) = 1 - q(IFS(j),IFS(0)) \quad (19)$$

It can be seen from equation (18) and (19) that $S \in (0,1)$. If the intuitionistic fuzzy set $S(IFS(j),IFS(0))$ is bigger, the Pareto solution intuitionistic fuzzy set is more similar to reference solution intuitionistic fuzzy set. If the Pareto solution is closer to the reference solution, the quality of Pareto solution is better.

In order to solve the high dimensional multi-objective optimization problem of displacement flow shop scheduling, this paper combines the intuitionistic fuzzy sets similarity with GA to solve this problem. According to the description of the displacement flow shop scheduling problem, the machining path of the workpiece is the same as the machining order of the workpiece on all machines [16-19]. Therefore, this paper adopts the integer to encode the scheduling process, and each chromosome is corresponding to an arrangement of n workpieces. For example, a chromosome $x = \{x_{i1}, x_{i2}, \cdots, x_{in}\}$, where $1 < i < n+1$ represents the workpiece number. Occurrence frequency of each i is 1, so that each position component of chromosome corresponds to an workpiece number. Continuous transformation of GA algorithm from real numbers to discrete processes is realized by ROV(Rank Order Value).

The specific steps of the proposed algorithm are as follows:

Step 1. Population initialization.

Initializing the parameters, NP genes X_j^{gen} are randomly generated to form the initial parent population, where gen is the evolution number of the current population. Creating an initial external file.

Step 2. Establishing an intuitionistic fuzzy set of reference solutions.

It solves the optimization problem by formula (1)-(4) based on GA. Constructing the reference solution value Y_0. Y_0 is mapped to the reference solution intuitionistic solution fuzzy set IFS(0).

Step3. Establishing Pareto solution intuitionistic fuzzy set.

Objective function values of each gene are calculated to get Pareto frontier Y_j^{gen}. Y_j^{gen} is mapped to Pareto solution intuitive fuzzy set $IFS^{gen}(j)$.

Step 4. Calculating the similarity between the intuitionistic fuzzy sets.

The similarity $S(IFS(j), IFS(0))$ between $IFS^{gen}(j)$ and IFS(0) of each gene X_j^{gen} in gen-th population is calculated according to equation (19), which will be regarded as the fitness value of GA algorithm to guide the algorithm evolution.

Step 5. Selection, crossover and mutation operations.

The binary championship method is used to select the best genes. A new gene (Pareto solution) is generated by partial mapping crossover (PMX) and SWAP mutation (SWAP), and a new sub-population is formed.

Step 6. Maintaining and updating external files.

Merging the sub-population and external files. Calculating their crowding distance, keeping Pareto solutions with large crowding distance, deleting Pareto solutions with small crowding distance and trimming them to obtain updated external files.

Step 7. Judging the termination condition.

Judging whether the fitness value of the group does not change continuously for a certain time or whether it meets the maximum iteration number. If No, gen=gen+1, return to Step 3. Otherwise, stop.

4. Experiment and analysis

In order to verify the superiority of the proposed algorithm and the feasibility, effectiveness of solving the multi-dimension targeted poverty alleviation data scheduling problem, six latest CEC standard test sets and ten targeted poverty alleviation test examples with different sizes are selected for experiment testing in this paper. We also compare our method with MDWW [20], ECOTD [21], HGA [22].

4.1. Experiment parameter

For better comparison, the same parameters are set as follows: population size NP=20, external file size $W_{\max}$ =20, maximum evolutionary iteration number maxgen=100. GA crossover probability P_c = 0.9, mutation probability P_m = 0.1. The genetic algorithm is used to optimize the single target q=10. In proposed method, λ =0.9, non-membership function parameter α_i =1.2.

The parameters of the test function in the CEC standard test set are as follows. Six standard test functions: MaF1, MaF2, MaF4, MaF5, MaF10 and MaF11. The target function number M=4, and other parameters are as in reference [23]. Standard test function code can be downloaded in the website (http://www.cercia.ac.uk/news/cec2017maooc/).

The parameters of the flow workshop test examples are set as follows: the maximum transport volume in each batch of completed workpiece h=5, unit inventory cost $\beta_i = 1$, unit delay cost γ_i =2.

4.2. Performance evaluation index

To verify the performance of the proposed algorithm. In this paper, three performance evaluation indexes are adopted: space distance (SP), general distance (GD) and C index.

1) Diversity index.

Space distance (SP) is used to evaluate the distribution uniformity of solution set in the target space. The calculation formula is as follows:

$$SP = \left(\sum_{i=1}^{W_{\max}} (\bar{d} - d_i)^2 / (W_{\max} - 1)\right)^{0.5} \tag{20}$$

$$d_i = \min_j (| f_1^i - f_1^j | + | f_2^i - f_2^j | + | f_3^i - f_3^j | + | f_4^i - f_4^j |) \tag{21}$$

Where $i, j = 1, \cdots, W_{\max}$. $\bar{d}$ is the mean value of d_i . If the SP value is smaller, the solution set distribution is more uniform.

2) Convergence index.

General distance (GD) is used to evaluate the convergence of the algorithm. The formula is as follows:

$$GD = (\sum_{i=1}^{W_{\max}} d_i^2)^{0.5} / W_{\max} \tag{22}$$

Where d_i is the Euclidean distance between the i-th nondominant solution and the reference solution. The smaller GD value is, the closer Pareto solution is to reference solution, and the better convergence is.

3) Coverage index.

C index is used to measure the relative merits between two optimal solution sets $\bar{A}$ and $\bar{B}$ obtained by optimizing the same problem by two algorithms A and B. The calculation formula is as follows:

$$C(\bar{A}, \bar{B}) = \frac{|\{b \in B \mid \exists a \in A, s.t. a \prec b\}|}{|B|} \tag{23}$$

Where $C(\bar{A}, \bar{B})$ denotes the solution percentage in A that are not inferior to solutions in B. $C(\bar{A}, \bar{B}) \in [0,1]$. If the value of $C(\bar{A}, \bar{B})$ is bigger, the solution set A is better than B. So the algorithm $\bar{A}$ is better too.

4.3. Results of CEC data set

The six MaF standard test functions selected from the CEC set have different characteristics. Their theoretical fronts have different shapes, which can reflect the complex properties of practical optimization problems. Therefore, the performance of the proposed evolutionary high-dimensional multi-objective optimization method to solve practical optimization problems can be tested. The average value of evaluation indexes with different algorithms is shown in Table 1.

In terms of SP, in MaF2, MaF5 and MaF10, the SP value of proposed method is lower than that of the other three algorithms. In MaF1 and MaF4, the SP value of proposed method is greater than HGA, but less than the other two algorithms. In MaF11, the SP value of proposed method is greater than MDWW and HGA but less than ECOTD. It shows that, on the whole, Pareto solution set distribution obtained by proposed method is better than other algorithms. For GD, in MaF1, MaF2, MaF4, MaF5 and MaF11, the GD value of proposed method is less than that of the other three algorithms. In MaF10, the GD value of proposed method is greater than HGA but less than that of the other two algorithms. It shows that Pareto solution obtained by proposed method is closer to the reference solution, and its convergence is better than other three methods.

Table 1. Performance evaluation on CEC set with different algorithms

Instance	Method	SP	GD	Instance	Method	SP	GD
MaF1	MDWW	4.076	12.345	MaF5	MDWW	1.816	4.652
	ECOTD	0.346	0.597		ECOTD	0.998	2.829
	HGA	0.105	0.491		HGA	0.965	2.427
	Proposed	**0.199**	**0.395**		Proposed	**0.156**	**2.423**
MaF2	MDWW	0.135	0.157	MaF10	MDWW	0.082	0.105
	ECOTD	0.104	0.153		ECOTD	0.043	0.099
	HGA	0.054	0.166		HGA	0.028	0.052

	Proposed	**0.040**	**0.129**		Proposed	0.008	0.068
MaF4	MDWW	1752.775	2321.199	MaF11	MDWW	4.445	11.131
	ECOTD	1079.271	1177.894		ECOTD	0.062	15.832
	HGA	459.294	1602.693		HGA	1.934	11.987
	Proposed	**674.533**	**553.518**		Proposed	**1.345**	**10.111**

4.4. Results of ten data test sets

The results are the average values after 10 iterations as shown in Table 2. In instance 2, proposed method has 4 targets that are better than MDWW and ECOTD, and 2 targets that are better than HGA. In instances 1,3,4,5,6,7,8,9,10, proposed method obtains the optimal multi-dimension solution better than the other three algorithms. This indicates that proposed multi-dimension optimization solution is superior to the three methods in solving the problem of multi-dimension replacement flow shop scheduling.

Table 2. Results of 10 permutation flow shop test instances

Instance number	Method	Problem(n×m)	Reference solution $Y_0 = (f_1 f_2 f_3 f_4 f_5)$	multi-dimension optimization solution $(f_1 f_2 f_3 f_4 f_5)$	S
1	MDWW	10×5	(836,246,965,1586)	(879,302,1134,2028)	0.89
	ECOTD			(909,572,1311,3031)	
	HGA			(999,543,1767,3677)	
	Proposed			**(911,314,1421,2052)**	
2	MDWW	10×10	(1037,184,839,1132)	(1122,207,1244,1299)	0.93
	ECOTD			(1230,502,1711,3364)	
	HGA			(1134,356,1138,2162)	
	Proposed			**(1051,234,979,1372)**	
3	MDWW	20×5	(1295,381,1679,1889)	(1428,466,2398,2654)	0.93

	ECOTD			(1511,840,2678,4436)	
	HGA			(1365,791,2091,4485)	
	Proposed			**(1324,463, 1827, 2299)**	
4	MDWW	20×10	(1813,812,2123,6905)	(2045,1498,3339,10869)	0.87
	ECOTD			(1986,1112,2726,10435)	
	HGA			(1943,1312,2725,10061)	
	Proposed			**(1881,978,2284,7717)**	
5	MDWW	40×10	(3031,2158,4412,55231)	(3372,3015,6262,70943)	0.91
	ECOTD			(3232,2684,5483,67922)	
	HGA			(3138,2460,5155,61649)	
	Proposed			**(3036,2281,4703,56987)**	
6	MDWW	40×20	(3805,2496,4729,39029)	(4168,3450,6868,52750)	0.88
	ECOTD			(4067,3150,5868,53798)	
	HGA			(4025,2917,5798,47696)	
	Proposed			**(3855,2598,5055,41800)**	
7	MDWW	50×20	(4439,3164,6048,81178)	(4880,4328,8312,100202)	0.91
	ECOTD			(4751,3729,7552,104394)	
	HGA			(4612,3514,7043,90562)	
	Proposed			**(4504,3339,6208,85398)**	
	MDWW			(5539,4888,9331,132431)	
8	ECOTD	60×20	(5181,3857,7342,103623)	(5483,4595,8866,121752)	0.89
	HGA			(5315,4296,8249,114985)	

	Proposed			**(5029,3963,7571,106259)**	
	MDWW			(6743,5538,12056,129987)	
	ECOTD			(6637,4915,11231,126056)	
9		80×20	(6347,4198,9544,104914)		0.89
	HGA			(6527,4908,9985,124889)	
	Proposed			**(6403,4538,9698,106572)**	
	MDWW			(7980,6967,14595,247182)	
	ECOTD			(7828,6374,13756,240925)	
10		100×20	(7472,5624,9791,206859)		0.91
	HGA			(7712,6174,13113,217208)	
	Proposed			**(7519,5869,11913,208313)**	

According to Table 2, the intuitive fuzzy set similarity (S) of the 10 instances are all greater than 0.8. The Pareto optimal solution obtained by proposed method is similar to the reference solution.

In terms of SP, the SP value of proposed method is lower than that of the other three algorithms. The results show that the distribution uniformity of Pareto solution set obtained by proposed method is better than other algorithms. In terms of GD, the GD value of proposed method is lower than that of the other three algorithms, indicating that Pareto solution obtained by proposed method is closer to the reference solution and its convergence is optimal.

The coverage indexes results with different algorithms are shown in Table 3. It can be seen from the table that the optimization effect of the proposed method is obviously better than other methods. It is concluded that the proposed method is effective in solving multi-dimension targeted poverty alleviation data scheduling problems and more suitable for solving large-scale problems.

Table 3. “C” comparison with different methods

Instance number	Problem(n×m)	MDWW	ECOTD	HGA	Proposed
1	10×5	1	0.42	0.45	0.57
2	10×10	1	0.35	0.37	0.46
3	20×5	1	0.33	0.35	0.27
4	20×10	1	0.32	0.36	0.18
5	40×10	1	0.38	0.37	0.085

6	40×20	1	0.22	0.18	0.051
7	50×20	1	0.21	0.16	0.051
8	60×20	1	0.21	0.14	0.022
9	80×20	1	0.25	0.17	0.021
10	100×20	1	0.26	0.13	0.011

5. Conclusion

In this paper, we propose a new method based on the similarity of intuitionistic fuzzy sets to solve the high-dimensional multi-dimension targeted poverty alleviation data scheduling problem. The reference solution and Pareto solution are mapped to the intuitionistic fuzzy set respectively. The similarity of Pareto solution intuitionistic fuzzy set and Pareto solution is calculated and used as the fitness value of the evolutionary algorithm. A genetic algorithm based on the similarity of intuitionistic fuzzy sets is established to solve the multi-dimension targeted poverty alleviation data scheduling problem of four targets. The simulation results show that the similarity of intuitionistic fuzzy sets can be effectively combined with the genetic algorithm, and the performance of the algorithm is better than that of the common multi-objective optimization algorithm. Moreover, a high-quality Pareto solution can be obtained in the multi-objective replacement flow shop scheduling problem, which is more suitable for solving large-scale problems. In this paper, the multi-dimension targeted poverty alleviation data scheduling problem is selected to solve the shop scheduling problem. In actual production, more complex situations may occur, such as zero idle flow shop scheduling, blocked flow shop scheduling, job shop scheduling, flexible job shop scheduling, dynamic scheduling, etc. The next step is to explore the above problems by improving the algorithm, and further expand the application scope of the algorithm proposed in this paper.

Acknowledgements.

The author would like to thank the anonymous reviewers for their comments.

References

[1] Ta Q C, Billaut J C, Bouquard J L. Matheuristic algorithms for minimizing total tardiness in the m-machine flow-shop scheduling problem. Journal of Intelligent Manufacturing, 29:617-628, 2018.

[2] Wang D Z, Liu S X, Guo X W. A Multi-agent Evolutionary Algorithm for Solving Total Tardiness Permutation Flow-shop Scheduling Problem. Acta Automatica Sinica, 40(3):548-555, 2014.

[3] Zheng T, Yamashiro M. Solving flow shop scheduling problems by quantum differential evolutionary algorithm. International Journal of Advanced Manufacturing Technology, 49(5-8):643-662, 2010.

[4] Lin J. A hybrid discrete biogeography-based optimization for the permutation flowshop scheduling problem. International Journal of Production Research, 54(15-16):1-10, 2017.

[5] Gang L, Chong-Zheng H. Permutation Flow Shop Scheduling Based on Hybrid Discrete Artificial Bee Colony Algorithm. Control Engineering of China, 24(9): 1925-1929, 2017.

[6] Lu C, Gao L, Li X, et al. Energy-efficient permutation flow shop scheduling problem using a hybrid multi-objective backtracking search algorithm. Journal of Cleaner Production, 144:228-238, 2017.

[7] B. Li and L. Wang. A Hybrid Quantum-Inspired Genetic Algorithm for Multiobjective Flow Shop Scheduling. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 37(3): 576-591, 2007.

[8] Li H, Yin S, Liu J, et al. Novel gaussian approximate filter method for stochastic non-linear system. International Journal of Innovative Computing, Information and Control. 13(1): 201-218, 2017.

[9] Shoulin Yin, Ye Zhang, Shahid Karim. Large Scale Remote Sensing Image Segmentation Based on Fuzzy Region Competition and Gaussian Mixture Model. IEEE Access. 6:26069-26080. 2018.

[10] Teng Lin, Hang Li and Shoulin Yin. Modified Pyramid Dual Tree Direction Filter-based Image De-noising via Curvature Scale and Non-local mean multi-Grade remnant multi-Grade Remnant Filter. International Journal of Communication Systems. 31(16), 2018.

[11] P. Li, Z. Chen, J. Gao, J. Zhang, S. Jin, W. Zhao, F. Xia, and L. Wang, A Deep Fusion Gaussian Mixture Model for Multiview Land Data Clustering, Wireless Communications and Mobile Computing, vol. 2020, Article ID 8880430, 9 pages, 2020.

[12] Ananthi,V.P., Balasubramaniam,P. & Raveendran, P. A thresholding method based on interval-valued intuitionistic fuzzy sets: an application to image segmentation. Pattern Anal Applic 21, 1039–1051 (2018). https://doi.org/10.1007/s10044-017-0622-y

[13] Zne-Jung Lee, Shun-Feng Su and Chou-Yuan Lee. Efficiently solving general weapon-target assignment problem by genetic algorithms with greedy eugenics. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 33(1): 113-121, 2003.

[14] J. Gao, P. Li, and Z. Chen. A Canonical Polyadic Deep Convolutional Computation Model for Big Data Feature Learning in Internet of Things. Future Generation Computer Systems, 2019,99: 508-516.

[15] J Dhivya, B Sridevi. A novel similarity measure between intuitionistic fuzzy sets based on the mid points of transformed triangular fuzzy numbers with applications to pattern recognition and medical diagnosis. Applied Mathematics - A Journal of Chinese Universities, 34: 229–252(2019).

[16] Shoulin Yin, Hang Li, Asif Ali Laghari, et al. A Bagging Strategy-Based Kernel Extreme Learning Machine for Complex Network Intrusion Detection[J]. EAI Endorsed Transactions on Scalable Information Systems. http://dx.doi.org/10.4108/eai.6-10-2021.171247

[17] Qingwu Shi, Shoulin Yin, Kun Wang, Lin Teng and Hang Li. Multichannel convolutional neural network-based fuzzy active contour model for medical image segmentation. Evolving Systems (2021). https://doi.org/10.1007/s12530-021-09392-3

[18] Shoulin Yin, Hang Li, Desheng Liu and Shahid Karim. Active Contour Modal Based on Density-oriented BIRCH Clustering Method for Medical Image Segmentation [J]. Multimedia Tools and Applications. Vol. 79, pp. 31049-31068, 2020.

[19] S. Yin and H. Li. Hot Region Selection Based on Selective Search and Modified Fuzzy C-Means in Remote Sensing Images[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 13, pp. 5862-5871, 2020, doi: 10.1109/JSTARS.2020.3025582.

[20] Shao Z, Pi D, Shao W. A novel multi-objective discrete water wave optimization for solving multi-objective blocking flow-shop scheduling problem. Knowledge-Based Systems, 165(FEB.1):110-131, 2019.

[21] Binghai Zhou, Wenlong Liu. Multi-Objective Hybrid Flow-Shop Scheduling Problem Considering Energy Consumption and On-Time Delivery. Journal of Shanghai Jiaotong University, 053(007):773-779, 2019.

[22] Huang X, Yang L. A hybrid genetic algorithm for multi-objective flexible job shop scheduling problem considering transportation time. International Journal of Intelligent Computing and Cybernetics, 12(2):154-174, 2019.

[23] Cheng R, Li M, Tian Y, et al. A benchmark test suite for evolutionary many-objective optimization. Complex & Intelligent Systems, 3(1):67-81, 2017.