

A New Maximum Entropy Estimation of Distribution Algorithm to Solve Uncertain Information Job-shop Scheduling Problem

Lu Lin^{1,2}

¹School of Business Administration, Guizhou College of Finance and Economy, Guiyang China

²Guizhou Key Laboratory of Economic Simulation, Guiyang China

E-mail: lulin355@163.com

Abstract—Estimation of Distribution Algorithm (EDA) is a new kinds of colony evolution algorithm, through counting excellent information of individuals of the present colony EDA construct probability distribution model, then sample the model produces new generation. To solve the NP-Hard question as EDA searching optimum network structure a new Maximum Entropy Distribution Algorithm (MEEDA) is provided. The algorithm takes Jaynes principle as the basis, makes use of the maximum entropy of random variables to estimate the minimum bias probability distribution of random variables, and then regard it as the evolution model of the algorithm, which produces the optimal/near optimal solution. Then this paper presents a rough programming model for job shop scheduling under uncertain information problem. The method overcomes the defects of traditional methods which need pre-set authorized characteristics or amount described attributes, designs multi-objective optimization mechanism and expands the application space of a rough set in the issue of job shop scheduling under uncertain information environment. Due to the complexity of the proposed model, traditional algorithms have low capability in producing a feasible solution. We use MEEDA in order to enable a definition of a solution within a reasonable amount of time. We assume that machine flexibility in processing operations to decrease the complexity of the proposed model. Muth and Thompson's benchmark problems tests are used to verify and validate the proposed rough programming model and its algorithm. The computational results obtained by MEEDA are compared with GA. The compared results prove the effectiveness of MEEDA in the job shop scheduling problem under uncertain information environment.

Index Terms—*job shop scheduling under uncertain information, rough programming, maximum entropy estimation of distribution algorithm, test*

I. INTRODUCTION

In the field of science, engineering and economy, many of latest the advances rely on computing corresponding optimization problems of the global optimal solution of the numerical technique.

EDA (Estimation of Distribution Algorithm) is a new algorithm based on the population evolution, it constructs the probability distribution model by statistic the information of optimal individuals in a present group, then samples the model to produce the next generation, thus, the offspring can be distributed in good space with large probability, which can promote the group's evolution. Compared with the traditional evolutionary algorithm, crossover and mutation operator is replaced by distribution estimation, what else the new population produced by distribution estimation. EDA makes use of the distribution of random vectors to characterize its internal probability of dependence, therefore how to find an n-dimensional probability model which can fully reflect the interaction between random variables is the key of EDA. Probability distribution model directly determines the performance of EDA. Precise probability distribution model ensures a good outcome, and vice versa. Currently, the more common practice is the application of probability map model (Reference [1] shows that discrete EDA uses bayes network to express the relationship between variables, continuous EDA uses gauss network). However, how to determine the network parameters which in turn determine the optimal network structure is proved to be a NP-Hard problem, and the complicated calculation has greatly reduced the computing performance of EDA and limited the application of EDA. This paper presents a new algorithm named Maximum Entropy Estimation of Distribution Algorithm (MEEDA), this algorithm is based on Jaynes principle, using the random variables' maximum entropy estimates its minimum bias probability distribution, and as the algorithm's evolution model, and efficiently reduce the computational complexity of the algorithm. The simulation experiments of the two kinds of problems prove the effectiveness of the proposed algorithm, with higher global search ability and more stable convergence.

II. MAXIMUM ENTROPY ESTIMATION OF DISTRIBUTION ALGORITHM

A. The principle of EDA

EDA is generated from genetic algorithm (GA). GA performs selection, producing a mating pool; a new population is then produced from the mating pool by crossover and mutation operations. This process is then iterated until a termination criterion is met. Although GA can be applied to many combinatorial optimization problems, yet they all need some prior knowledge to determine the evolution operator, therefore, it is difficult to find an appropriate set of parameters in coding and combined with the operation for a particular problem, which is the inherent defect of GA. Furthermore, in evolutionary terms, the crossover and mutation operator are used to the realization of and the exchange of individual information in group, and select operator is used to the guidance of the evolutionary process, both of them did not consider the relevant information of the same generation individuals, so they did not fully make use of the existing information and reduced the operating efficiency of GA; furthermore, crossover and mutation operator can not guarantee the establishment of building blocks assumptions in theory, as a result, GA demonstrates very poor performance for some of the issues (such as misleading).

To improve these deficiencies, H. Muhlenbein and G. Paa β in reference [2] proposed the estimation of distribution algorithm in 1996, which quickly aroused the concern of extensive research. And it has been a hotspot of evolutionary computation at present. EDA does not use crossover and mutation operator, but uses probability distribution of higher degree individuals in the group as the evolutionary model, and then produces the next generation of subgroups; with the use of tracking advanced method to replace the recombination of the GA modes, and it has expanded the applied space of EDA.

As an evolution model is derived from the statistical probability distribution of information, thus it shows the main characteristics of groups in making the best use of existing information and reflecting more accurately the relationship between variables. The theoretical study in reference [3] shows that in the iterative process, EDA may get interaction information of individuals between group and different bit of individuals, identify and manipulate the model of important blocks, therefore it can effectively solve the optimization problems in which the decision-making variables are interactive.

Reference [1] shows a schematic of the EDA approach. In the first step N individuals are generated at random in Figure 1, for example, based on an uniform distribution (discrete or continuous) on each variable These N individuals constitute the initial population, D_0 , and each of them is evaluated. In a second step, a number Se ($Se < N$) of individuals are selected

(usually those with the higher objective function value). Next, the induction of the n-dimensional probabilistic model that better reflects the interdependences between the n variables is carried out. In a fourth step, N new individuals (the new population) are obtained by means of the simulation of the probability distribution learnt in the previous step. The steps 2, 3 and 4 are repeated until a stopping condition is verified. Examples of stopping conditions are: when a fixed number of populations or a fixed number of different evaluated individuals are achieved, uniformity in the generated population, and not improvement with regard to the better individual obtained in the previous generation.

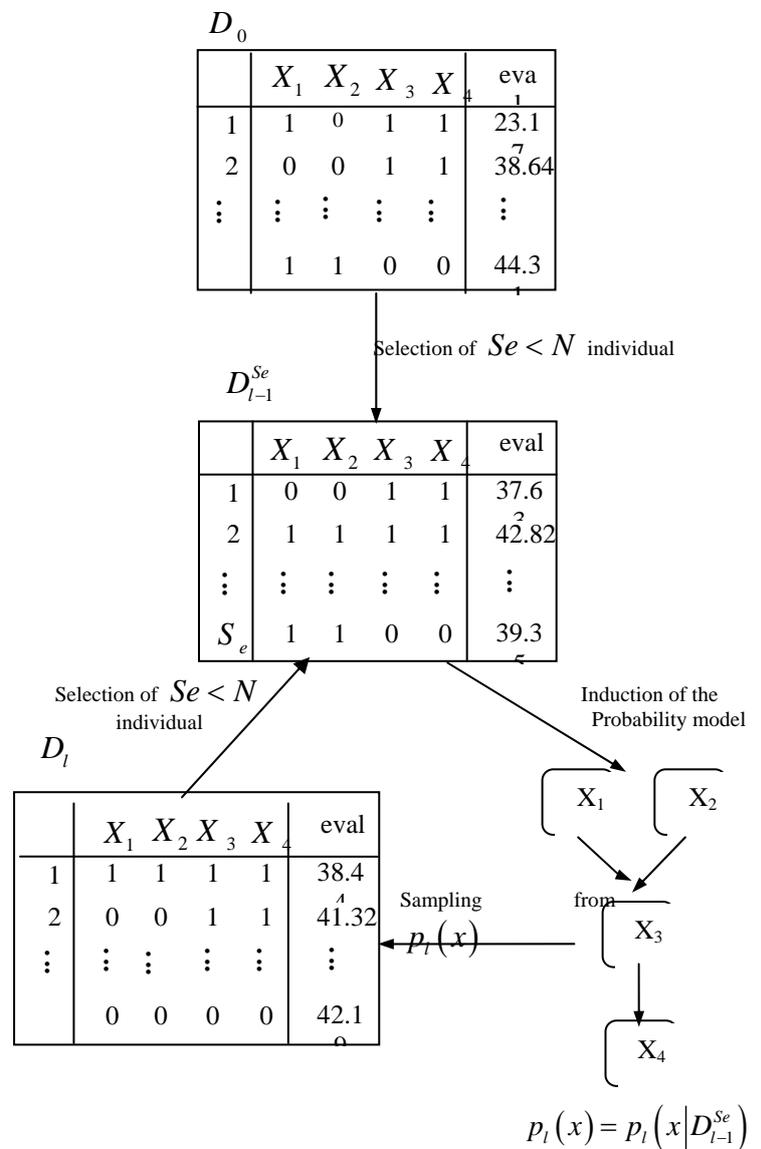


Figure 1: Illustration of the EDA approach to optimization

EDA makes use of the distribution of random vector to characterize its internal probability of dependence, therefore how to find a n -dimensional probability model which can fully reflect the interaction between random variables is the key of EDA. Probability distribution model directly determines the performance of EDA. Precise probability distribution model ensures a good outcome, and vice versa.

James R. Neil.al. show the pseudo code for an EDA. The factorization step takes the population after s election and constructs a graphical model. The graphical model is then sampled population size times using probabilistic logic sampling to produce the next population.

```

Population=random Population (population Size);
while (terminating conditions to met)
{
Gene pool = selection (population);
//Factorization step
bayesNet = buildBayesNetFromSample (genepool);
//End Factorisation Step
//Sampling step
For (I = 0; i<populationSize; i=i+1)
{
newPopulation[i] = PLS (bayesNet);
}
//End Sampling step
Population = newPopulation;
}
    
```

Good surveys of existing algorithms can be founding in Larranagaet.al.[4] and in Pelikan[5,6].

Currently, the more common practice is the application of probability map model (discrete EDA uses bayes network to express the relationship between variables, continuous EDA uses gauss network). Let's consider continuous random variables. Assumption of the edge probability distribution and the statistical probability distribution of x_i are normal distribution, based on Gauss network, probability density $P(x)$ can be said for the multiplication of each separate conditional probability which obeys normal distribution. In the equation of

$$P(x) = \prod_{i=1}^n P(x_i | pa(x_i)) = \prod_{i=1}^n N\left(\mu_i + \sum_{j=1}^{i-1} \beta_{ij}(x_j - \mu_j), \sigma_i\right) \quad (1)$$

μ_i is the unconditional mean of variable x_i , σ_i is the variance in the condition of given variable x_1, \dots, x_{j-1} , β_{ij} is a linear coefficient of the relationship intensity between x_i and x_j . From the multiplicative theorem of probability, there are $n!$ networks has this joint probability distribution. How to acquire network parameters to determine the optimal network structure is proved to be an NP-Hard problem, and the complicated calculation has greatly reduced the computing performance of EDA and limited the application of EDA. The existing general method is the use of greedy method search network architecture, or solves the Gauss network parameters through the covariance matrix. These

methods cannot guarantee to found the most superior network architecture, or reduce the algorithm operating speed. Considered in the theory of probability, we can use the entropy to express system's average uncertainty, but the random event uncertainty size may be described by the probability distribution function, like this, random variable's biggest entropy distribution is the random variable's smallest prejudice probability distribution, namely most may several distribute, thus, simultaneously avoided the complex computation which for the simple and direct construction precise probabilistic model sought the design parameter to carry on providing the possibility.

This paper presents a new algorithm named Maximum Entropy Estimation of Distribution Algorithm (MEEDA), this algorithm is based on Jaynes principle, using the random variables' maximum entropy estimate its minimum bias probability distribution, and as the algorithm's evolution model, efficiently reduce the computational complexity of the algorithm.

B. The MEEDA principle

Maximum entropy method estimates the probability distribution of random variables on the basis of the entropy of random variables, the theory of which is based on the principle of Jaynes[7], that is "the most realistic probability of distribution gives the maximum entropy of bound information". To the continuous random variables, entropy can be definition from the following equation:

$$H(x) = -\int_R P(x) \log(x) dx$$

(2)

In the equation, $P(x)$ is the function of probability density of random variable X , R is the definition domain of $P(x)$. Maximum entropy distribution gets the maximum entropy $H(x)$ by adjusting $P(x)$ on the conditions of statistical sample. According to the varying points, the author uses Lagrange indefinite multipliers to act as functional objectives:

$$L = \int_R [-P(x) \ln P(x) + (\lambda_0 + 1)P(x) + \sum_{i=1}^m \lambda_i x^i P(x)] dx \quad (3)$$

The density function is:

$$P(x) = \exp(\lambda_0 + \sum_{i=1}^m \lambda_i x^i) \quad (4)$$

$P(x)$ can be fully identified by further defining parameters λ . Through the derivation of the mathematics, λ_0 is calculated as follows:

$$\lambda_0 = -\ln\left[\int_R \exp\left(\sum_{i=1}^m \lambda_i x^i\right) dx\right] \quad (5)$$

λ_i ($i = 1, 2, \dots, m$) can be obtained from value(6):

$$\delta_i = 1 - \frac{\int_R x^i \exp(\sum_{i=1}^m \lambda_i x^i) dx}{\mu_i \int_R \exp(\sum_{i=1}^m \lambda_i x^i) dx} \quad (6)$$

The approximate solution of λ_i can be got by calculating the minimum $\min \delta_i^2$ of square of residuals δ_i of value(6).

C. The calculation steps of MEEDA

Step1: randomly generate N initial solutions; calculate initial fitness of these initial solutions.

Step2: choose S_e individuals which have higher adaptation degree ($S_e \leq N$).

Step3: take these S_e individuals values into the equation (5), (6), calculate values of parameters λ_0 and λ_i .

Step4: calculate the probability density function $P(x)$ of these S_e individuals under equation (4).

Step5: calculate the maximum entropy distribution function of these S_e individuals under equation (2), establish optimal n victories probability model that reflects the interaction of random variables.

Step6: produce N new entity by n victories probability model.

Step7: calculate the adapt degree of these new individuals, and replace the parent populations with new ones.

Step8: if the termination of the conditions is met, you can put out the result; otherwise you can turn to Step2.

D. The performance analysis of MEEDA

The main calculation of MEEDA focuses on the set of network structure. At present, the computing amount of the common use of greedy to search network structure is $o(n^2)$; through covariance matrix to get gauss network parameters is $o(n^3)$ computing amount. They are index, only suitable for small networks. From the above MEEDA, the computing amount of steps3 is $o(n \log n)$, steps4 is $o(n)$, steps5 is $o(n)$. Therefore, the time complexity of MEEDA is polynomial which is $o(n \log n + n + n)$, compared to the computing complexity of $o(n^2)$ of greedy and $o(n^3)$ of covariance matrix method, it has greatly reduced the computation amount.

III THE MATHEMATICAL MODEL OF JSSP UNDER UNCERTAIN INFORMATION

A. Job shop scheduling problem

Job shop scheduling problem is one of the most significant issues in production planning. Job shop

scheduling is significant because it determines process maps and process capabilities for most industries. Job shop scheduling problem of workshop production process is the core of manufacturing systems operations technology, management technology and the development of optimization technology. Relevant data in reference [8] shows that 95 percent of the manufacturing process time is consumed in the non-cutting process; it showed that the research and application of effective method of scheduling and optimization technology have become the basis and the key of advanced manufacturing technology practice. Job shop scheduling problem (JSSP) is a kind of task allocation which meeting the requirements of the order and being constrained to the allocation of resources, in essence, it is three-dimensional multi-objective decision-making problem involving task, time and resources. In a job shop scheduling problem with n jobs and m machines, there is $(n_1)!(n_2)!(n_3)! \cdots (n_m)!$ sequences.

n_k is the number of operations that must be executed by machine k . JSSP has been proven as NP-complete problem in the most difficult issues. This is a central issue in production management and combinatorial optimization. In the last three decades, many researchers have become interested in such problems. Certain JSSP has been well studied, the general assumption is that the processing time of work piece in specific equipment is known and do not consider the difference of work piece in priority. JSSP under uncertain information concerning to the uncertainty parameters (for example, in the process of the handling capacity of products, the processing time and consumption costs, etc.). Taking into account the important of flexible manufacturing in modern manufacturing industry, such problem has obvious practical significance.

Jia and Ierapetritou[9] handled uncertainty in short-term scheduling based on the idea of inference-based sensitivity analysis for MILP problems and utilization of a branch-and-bound solution methodology. The basic idea of the proposed method is to utilize the information obtained from the sensitivity analysis of the deterministic solution to determine (a) the importance of different parameters and constraints; (b) the range of parameters where the optimal solution remains unchanged. The main steps of the proposed approach are shown in Fig. 2. Jia and Ierapetritou[10] proposed a new method of uncertainty analysis on the right hand side (RHS) for MILP problems. The proposed solution procedure starts with the B&B tree of the MILP problem at the nominal value of the uncertain parameter and requires two iterative steps, Linear Programming (LP)/multi-parametric Linear Programming (mpLP) sensitivity analysis and updating the B&B tree (Fig. 3).

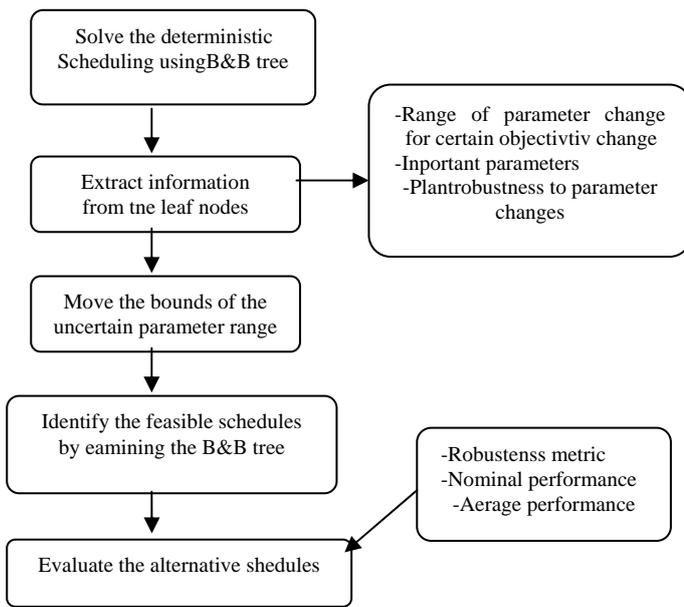


Figure 2. Flowchart of the sensitivity analysis method

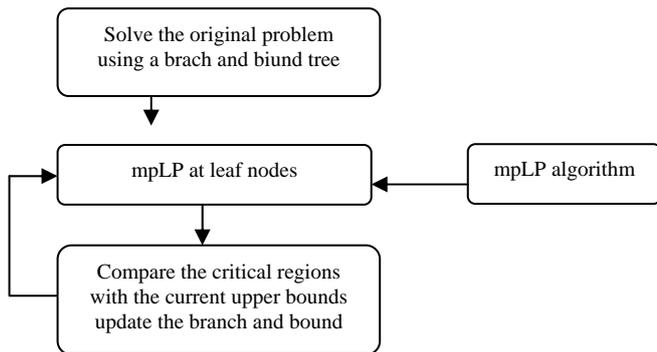


Figure 3. Flow chart of uncertainty analysis

At present, there are four basic methods to dealing with the uncertain information: (1) random probability theory; (2) fuzzy set theory; (3) Gray theory; and (4) rough set theory. Zhang Guo-jun.al.[11]used random variables to describe uncertainty Job-shop parameters, developed a hybrid intelligent algorithm which fuses random simulation, neural network and genetic algorithm to solve such uncertain scheduling model; Xu Zhen-hao.al.[12] constructed a scheduling model which used triangle fuzzy numbers to process uncertainty processing time and proposed immune scheduling algorithm of such problem; Li Bing, Gu Xing-sheng [13]studied grey chance constrained programming for job shop scheduling under uncertainty, and gives the corresponding genetic algorithm. The use of these methods has a theoretical premise that the uncertain information according to the random distribution or fuzzy function or whitening weight function. Rough set is a new mathematical tool dealing with ambiguity and inaccuracy problem

proposed by Polish scientist Zdzislaw Pawlak [14] in the early 1980s without the pre-default attributes or the number description of characteristics such as the statistical probability distribution, fuzzy set theory of membership degree or membership function, etc, it identifies inherent law in approximate domain of problems which be determined from undistinguished relations and undistinguished types of the given set. Yu Ai-qing, Gu Xing-sheng [15] presented rough scheduling GA for flow shop scheduling problem under uncertainty. This paper constructs rough programming model of JSSP under uncertainty based on the constrained of uncertain processing time of work piece, and proposes a new maximum entropy estimation of distribution algorithm (MEEDA) to solve such complex problems.

It is assumption that a processing system has $N\{n_1, n_2, \dots, n_N\}$ machines and $N\{n_1, n_2, \dots, n_N\}$ work pieces, each part includes $K_i (K_i = 1, 2, \dots, M)$ procedure, to remember $J_i(o_{i1}, o_{i2}, \dots, o_{iK_i})$; the process sequence of work piece is pre-set, every procedure O_{iK_i} can be processed on multiple different machines lathe, the processing time of procedure O_{iK_i} is changed according to different performance of machines. The scheduling goal is to choose the most appropriate machinery for each procedure to identify the best order and the best start time of every work piece procedure in each machine so as to achieve certain optimal performance.

B. Rough set theory

The main ideology of rough set is to partite the domains based on currently available knowledge of given issue, while maintaining the same classification ability, it derives the concept of classification rules through knowledge reduction.

Given the non-vacuous finite set U , known as the domain; R is equivalent to the relationship of U , known as undistinguished relations, it divides the elements of U into various intersection subsets, then formats equivalence classes; $A = (U, R)$ is known as the approximate space. The general definition of rough set using two concepts: upper approximations and lower approximations. Approximate set is that when a collection can not be properly classified by effectively using equivalence relations, the other set can be adopted to achieve this similar collection. Subset $X \subseteq U$ can be described by two basic operations:

Lower approximations:

$$\underline{A}(X) = \bigcup \{Y \subseteq U \mid R : Y \subseteq X\} \quad (7)$$

Upper approximations:

$$\bar{A}(X) = \cup \{Y \subseteq U \mid R : Y \cap X \neq \emptyset\} \quad (8)$$

Let R be a real set, (a, b) is on the open range of R , the real number sequence $S = \{x_0, x_1, \dots, x_n\}$ of (a, b) makes $a = x_0 < x_1 < \dots < x_n = b$, then $A = (R, S)$ is known as approaching space the generated by S , S is known as discrete series, each S has breakdown definition on (a, b) :

$$\pi(S) = \{\{x_0\}, (x_0, x_1), \{x_1\}, (x_1, x_2), \dots, (x_{n-1}, x_n), \{x_n\}\}$$

This equivalent to a equivalence definition, so that for $\forall x \in (a, b)$, its under approximation and on approximation of S can be defined as:

$$apr_S(x) = \sup \{x_i \in S : x_i \leq x, i = 0, 1, \dots, n\} \quad (9)$$

$$9) \quad \bar{apr}_S(x) = \inf \{x_i \in S : x_i \geq x, i = 0, 1, \dots, n\} \quad (10)$$

Clearly, when $x = x_i$, $apr(x) = \bar{apr}(x) = x_i$, while $x_i < x < x_{i+1}$, $apr(x) = x_i$, $\bar{apr}(x) = x_{i+1}$, were the forward-backward points of equivalence classes.

C. Mathematical models

Let rough variables $r = \left(\begin{smallmatrix} \bar{r} \\ r \end{smallmatrix} \right)$ describe the uncertainty of the processing time of work piece, r is the lower approximations describes the minimum determination time of the processing time of work piece, \bar{r} is the upper approximations describes the most likely time of the processing time of work piece.

Let rough variable t_{m,m_j} be the processing time of work piece $n_i (i = 1, 2, \dots, N)$ on machine $m_j (j = 1, 2, \dots, M)$, rough variable p_{m,m_j} be the completion time of processing of work piece $n_i (i = 1, 2, \dots, N)$ on machine $m_j (j = 1, 2, \dots, M)$. Then the completion time of N work pieces and M machines of JSSP under uncertain time is:

$$p_r(n_1, 1) = t_{m_1, 1} \quad (11)$$

$$p_r(n_1, m_j) = p_r(n_1, k-1) + t_{m_1, m_j} \quad j = 2, 3, \dots, M \quad (12)$$

$$p_r(n_i, 1) = p_r(n_{i-1}, 1) + t_{m_i, 1} \quad i = 2, 3, \dots, N \quad (13)$$

$$p_r(n_i, m_j) = \max \{p_r(n_{i-1}, m_j), p_r(n_i, m_{j-1})\} + t_{m_i, m_j} \quad i = 2, 3, \dots, N \quad j = 2, 3, \dots, M \quad (14)$$

Let $H(t)$ be the completion time of the last work piece of $p_{\max} = p_r(j_N, m)$ on the final machine, then the objective function is $\min H(t)$.

D. Model solution

For the rough variables of the interval number, according to some theorems and expansion definition of rough set, the paper definite rough operations and concept for JSSP under uncertain information.

① Rough calculation is rough addition, the definition of two rough variables $x = \left(\begin{smallmatrix} \bar{x} \\ x \end{smallmatrix} \right)$ and $y = \left(\begin{smallmatrix} \bar{y} \\ y \end{smallmatrix} \right)$, rough addition is defined as:

$$x + y = \left(\begin{smallmatrix} \bar{x} \\ x \end{smallmatrix} \right) + \left(\begin{smallmatrix} \bar{y} \\ y \end{smallmatrix} \right) = \left(\begin{smallmatrix} \bar{x} + \bar{y} \\ x + y \end{smallmatrix} \right) \quad (15)$$

② The information described by rough variables is inaccurate or the uncertainty, therefore, definite *precision* as imprecise measure of rough variables:

$$precision(r) = \frac{\bar{r} - r}{Range(U)} \quad (16)$$

$$\min o = \min(\alpha \times precision + \beta \times ration) \quad (17)$$

$Range(U)$ is the interval range of the domain of rough variable r .

According to the reference [16], the objective function of rough programming scheduling model can be transformed as:

$$\min o = \min(\alpha \times precision + \beta \times ration) \quad (18)$$

Among them, $ration = \frac{p_r(n_N, M)}{\max \left\{ p_r(n_N, M) \right\}}$;

α, β is the definition of parameters for decision-making, α describes the importance of *precision* on the goal of decision-making, the smaller *precision*, the higher precision of the rough variables. Similar to this, β describes the importance of *ration* on the goal of

decision-making, the smaller *ration*, the smaller lower approximations of *makespan*. α And β are determined based on the objectives demand of decision-making.

IV. SIMULATION

Three well-known standard test question MT06, MT10 and the MT20 proposed by Muth and Thompson [17] are used for MEEDA test. The technique routes of various work pieces are fixed, every procedure can process in multiple devices, only differences on the processing time and cost. Table I gives the raw data of MT06, (m,t) describes work piece corresponding procedures on machine m and processing time t . In the actual operation of this paper, the lower approximations of work piece processing time t comes from the standards data, on the basis, the upper approximations is floating product according to actual production situation.

The set of parameters and operator strategy of MEEDA are as follows: $N=1000$, $S_e=500$, the highest adaptation individual in S_e can directly get into the next offspring, to compute the remaining 499 individuals to construct the probability distribution model. The set of parameters of GA are as follows: $N=10000$, exchange probability $p_c=0.90$, mutation rate $p_m=0.01$, using proportion choice selection mechanism to reproduce and two points exchange strategy where the point is randomly generated. The initial group of two algorithms is randomly generated according to normal in the same way. The decision-making parameters of rough programming model are that $\alpha=0.30, \beta=0.80$. The algorithm program is constructed by using Visual C++ language on a PC, the calculation results in figure 4, figure 5 and table 2. Among them, figure 4 is the Gantt of MT06, figure 5 is the convergence curve of MEEDA of MT10, table 2 is the performance compared of MEEDA and GA.

TABLE I.

RAW DATA OF MT06

	m,t	m,t	m,t	m,t	m,t	m,t
Job 1	3,1	1,3	2,6	4,7	6,3	5,6
Job 2	2,8	3,5	5,10	6,10	1,10	4,4
Job 3	3,5	4,4	6,5	1,9	2,1	5,7
Job 4	2,5	1,5	3,5	4,3	5,8	6,9
Job 5	3,9	2,3	5,5	6,4	1,3	4,1
Job 6	2,3	4,3	6,9	1,10	5,4	3,1

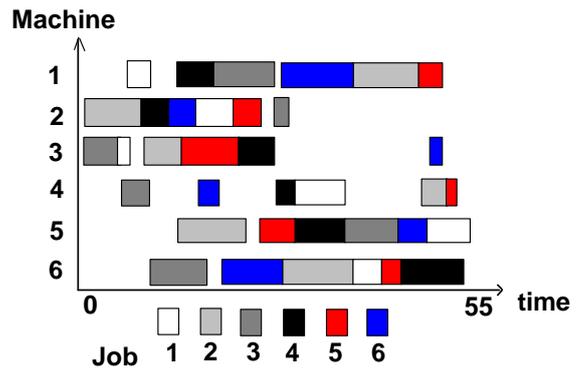


Figure 4. Gantt of MT06

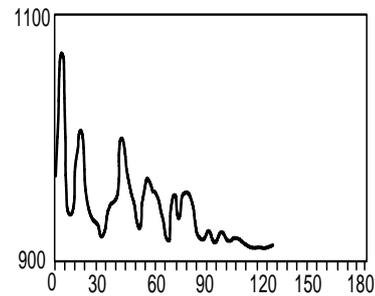


Figure 5. Convergence curve of MEEDA of MT10

TABLE II.

THE PERFORMANCE COMPARED OF MEEDA AND GA

problem	n	m	best solution	GA	MEEDA
MT06	6	6	55	55	55
MT10	10	10	930	965	930
MT20	20	5	1165	1215	1165

VI. CONCLUSION

This paper studies job shop scheduling problems under uncertain information, uses rough set theory to establish rough programming scheduling model of such problem, which is more truly reflects the uncertain situation in workshop than other scheduling models, and proposes a new maximum entropy estimation of distribution algorithm to solve such complex problems. The simulation verifies the validity of rough programming model and MEEDA, which provides a new approach for job shop scheduling problem under uncertain information.

ACKNOWLEDGMENT

The author gratefully acknowledges the talent introduction financial support from Guizhou College of Finance and Economy.

REFERENCES

- [1] Larrañaga, P., Etxeberria, R., Lozano, J. A., Peña, J. M.: Optimization by learning and simulation of Bayesian and Gaussian networks. Technical Report EHU-KZAA-IK-4/99, University of the Basque Country, December 1999.
- [2] Muhlenbein H, Paa β G. "From Recombination of Genes to the Estimation of Distribution Algorithms I. Binary Parameters". *Parallel Problem Solving from Nature*. Berlin: Springer, Vol. 4, 1996, pp.178-187.
- [3] Zhang Q, Mihieliebi H. "On the convergence of a class of estimation of distribution algorithms". *IEEE Trans on Evolutionary Computation*. Vol.8, No. 2, 2004, pp. 127-136.
- [4] J.A.LozanoP.Larraaga. Estimation of Distribution Algorithms.A New Tool for Evolutionary Computation.
- [5] Martin Pelikan,David E.Goldberg,and Fernando Lobo. A survey of optimization by building and using probabilistic models.IlliGAL Report No. 99018,Illinois Genetic Algorithms Laboratory,University of Illinois atUrbana-Champaign,Urbana, IL,1999.
- [6] M.Pelikan. Bayesian optimization algorithm:From single level to hierarchy.PhD thesis,University of Illinois,2002.
- [7] Siddall J.N. *Probabilistic engineering design*. Principles and Applications. New York: Marcel Dekker, 1983
- [8] He Ting , Liu Fei ,Ma Yulin; Yang Hai. "Study on shop-floor scheduling". *Chinese Journal of Mechanical Engineering*, Vol. 36, No.5, 2000,pp.97-102 (in Chinese).
- [9] Jia, Z., & Ierapetritou, M. G. . Short-term scheduling under uncertainty using MILP sensitivity analysis. *Industrial and Engineering Chemistry Research*, 43, 2004,pp.3782–3791.
- [10] Jia, Z., & Ierapetritou, M. G.. Uncertainty analysis on the right-handside for MILP problems. *AIChE Journal*, 52, 2006,pp.2486–2495.
- [11] Zhang Guo-jun, Li Chan-juan, Zhu Hai-pin,etc. "A Hybrid Intelligent Algorithm for Job-shop Scheduling under Uncertain Information Environment". *Chinese Mechanical Engineering*, Vol. 18 , No. 16, 2007,pp.1939-1942 (in Chinese).
- [12] Xu Zhen-hao, GU Xing-shen. "Immune scheduling algorithm for flow shop problems under uncertain". *Journal of Systems Engineering*, Vol. 20, No. 4, 2005, pp.374-380 (in Chinese).
- [13] Li Bing, Gu Xing-sheng. "Grey chance constrained programming for job shop scheduling under uncertainty". *Chinese High Technology Letters*, Vol. 16, No. 10, 2006,pp.1025-1029 (in Chinese).
- [14] Pawlak Z. "Rough sets". *International journal of computer and information sciences (S0091-7036)*, Vol. 11, No. 5, 1982, pp. 341-356.
- [15] Yu Ai-qing , Gu Xing-sheng. "Flow Shop Scheduling Problem under Uncertainty Based on Generalized Rough Sets". *Journal of System Simulation*, Vol. 18, No.12, 2006,pp.3369-3372, 3376 (in Chinese).
- [16] Lingras P. "Unsupervised Rough Set Classification Using Gas". *Journal of Intelligent Information Systems (S0925-9902)*, Vol. 16, 2001, pp.215-228.
- [17] Taillard E. "Benchmark for Basic Scheduling Problems". *European Journal of Operational Research (S0377-2217)*, Vol. 64, 1993, pp. 278-285.

Lu Lin is doctor of management at the Guizhou College of Finance and Economy. His key research interests include processes for logistics management and optimization calculation.