A new hybrid genetic algorithm for job shop scheduling problem

Ren Qing-dao-er-ji a, Yuping Wang b,*

a School of Science, Xi’an University, Xi’an 710071, China
b School of Computer Science and Technology, Xi’an University, Xi’an 710071, China

1. Introduction

Scheduling is one of the most critical issues in the planning and managing of manufacturing processes. The difficulty of finding the optimal schedule depends on the shop environment, the process constraints and the performance indicator. One of the most difficult problems in this area is the job shop scheduling problem (JSSP), which has been proved to be an NP-complete problem [1]. Since the benchmarks of JSSP were presented by Fisher and Thomson in 1963 [2], JSSP has been studied by a great number of researchers, and many exact methods and approximation algorithms have been proposed [3–7]. Exact methods, such as branch and bound, linear programming and Lagrangian relaxation guarantee global convergence and have been successful in solving small instances. However, they require a very high computing time as the size of problem increases. So a lot of researchers paid their attention to meta-heuristic and intelligent hybrid search optimization strategies for solving JSSP. Among these, the shifting bottleneck approach, particle swarm optimization, ant colony optimization, simulated annealing, Tabu search, genetic algorithm, neural network, immune algorithm are the typical examples. These intelligent optimization algorithms are relatively easy to implement and could conveniently be adapted to different kinds of scheduling problems, and especially, it has been proven by experiments that they can find high-quality solutions within reasonable computational time. These have made the research on intelligent hybrid algorithm for JSSP increasingly popular in the recent years.

Genetic algorithms were proposed by Holland [8] and have been successfully used in a variety of practical problems. Davis [9] first applied a genetic algorithm to the JSSP in 1985 successfully, and now genetic algorithms have been proved to be an effective approach for the JSSP. There are many such works, e.g., a genetic algorithm with search area adaptation was proposed by Someya and Yamamura [10]. Zhou and Feng [11] proposed a hybrid heuristics GA for JSSP, where the scheduling rules, such as shortest processing time (SPT) and most work remaining (MWKR), were integrated into the process of genetic evolution. Park et al. developed an efficient method based on genetic algorithm to address JSSP. The scheduling method based on parallel genetic algorithm was designed in [12]. Mattfeld and Bierwirth [13] considered a multi-objective job shop scheduling problem with release and due dates, as well as tardiness as objectives. Watanabe et al. [14] proposed a modified genetic algorithm with search area adaptation for solving the job shop scheduling problem. Li presented a two-row chromosome structured new genetic algorithm based on working procedure and machine distribution [15]. The relevant crossover and mutation operations were also designed in [15].

However, the existing genetic algorithms for the JSSP are usually with a slow convergence speed and easy to trap into local optimal solutions. In order to enhance the convergence speed, many researchers focused their attention on combining the genetic algorithm with local search schemes to develop some hybrid optimization strategies for JSSP. Wang and Zheng [16] developed a hybrid optimization strategy for JSSP by combining the simulated annealing algorithm and the genetic algorithm.
Gonçalves et al. [17] presented a hybrid genetic algorithm for the JSSP, in which the schedules were constructed by using a priority rule and the priorities were defined by the genetic algorithm, and then a local search heuristic was applied to improve the solution. Zhang proposed a genetic simulated annealing algorithm to solve the JSSP by combining the GA and SA [18]. Liu et al. [19] combined the GA and SA [18]. Zhang proposed a fast and elitist genetic algorithm based on NSGA-II for solving the multi-objective SSSP [21]. Zhang combined the TS algorithm and the genetic algorithm to develop a hybrid algorithm for JSSP [22]. Zhang proposed a hybrid simulated annealing algorithm based on a novel immune mechanism for the JSSP [23].

However, there are the following shortcomings for the aforementioned algorithms: (1) these algorithms have not taken into account the diversity of population. This can easily lead to the “premature” convergence because of the population being easily filled with more similar individuals. (2) Most of the crossover operators and mutation operators have not made use of the characteristics of the JSSP structure itself. Many researchers designed these operators according to the structure of code and only changed the form of code. Thus, these operators usually cannot guide the search to move to the better solutions, and are also difficult to integrate the merit of the parent individuals. (3) The local search ability is not satisfactory.

In order to overcome these drawbacks of the algorithms, in this paper, a new crossover operator and mutation operator were designed by sufficiently making use of the information and structure of the JSSP, and these operators are efficient for JSSP. Moreover, in order to enhance the diversity of the population and avoid trapping into the local optimal solutions, a specifically designed mixed selection scheme is presented. Furthermore, to improve the speed of the algorithm, an efficient local search approach is proposed and integrated into the proposed genetic algorithm. The detailed contributions are as follows:

(1) Clear definitions of the similarity and the concentration were given. And then a mixed selection operator based on the fitness value and the concentration value was proposed. This operator increased the diversity of the population and could prevent the “premature” convergence to some extent.

(2) New kinds of crossover operator based on the machine and mutation operator based on the critical path were specifically designed accordingly to the graph model of JSSP. To calculate the critical path, we presented a new algorithm for finding it from schedule.

(3) A local search operator was designed so as to improve the local search ability of GA.

(4) Based on these genetic operators, we proposed a hybrid genetic algorithm (HGA) and proved its convergence. Finally, the efficiency of the proposed algorithm was verified by computer simulations on some typical scheduling problems.

The remainder of the paper is organized as follows. In Section 2, the models of the JSSP are set up. In Section 3, a hybrid genetic algorithm to the JSSP is presented and its convergence is proved. Section 4 presents the experimental results. The conclusions are made in Section 5.

2. Modeling the job shop scheduling problem

The JSSP with which we are concerned can be described as follows [24]: There are \( n \) different jobs to be processed on \( m \) different machines. Each job needs \( m \) operations and each operation needs to be processed without preemption for a fixed processing time on a given machine. There are several constraints on jobs and machines:

- A job can visit a machine once and only once.
- There are no precedence constraints among the operations of different jobs.
- Preemption of operations is not allowed.
- Each machine can process only one job at a time.
- Each job can be processed by only one machine at a time.
- Neither release times nor due dates are specified.

The problem is to find a schedule to minimize the makespan, that is, to minimize the time required to complete all jobs.

In order to clearly describe the proposed genetic operators, we briefly introduce the disjunctive graph theory model for the JSSP [25].

Given an instance of job shop scheduling problem, we can associate it with a disjunctive graph \( G=(N,AE) \), where \( N \) is a set of all operations (seen as vertices), \( N=\{O_1, O_2, \ldots, O_n\} \), \( j=1, 2, \ldots, n \). \( O_j \) represents the \( j \)th operation of the \( j \)th job. 0 and 1 are two virtual processes, respectively, and represent the beginning and end of the schedule. \( A \) is a set of edges which connect the adjacent operations of the same job. \( A=\{(O_i, O_j) | i=1, 2, \ldots, n \} \cup (O_{ij}, O_{i+1}) | i=1, 2, \ldots, n-1 \} \cup (O_{im}, 1) | i=1, 2, \ldots, n \}. \) It reflects the processing constraint between the different operations of the same job (indicated with directed solid line in Fig. 1). \( E \) is a set of non-directed edges which connect the operations processed on the same machine. Non-directed edge is the edge there are two possible directions (indicated with dashed line in Fig. 1). \( E=E_1 \cup E_2 \cup \ldots \cup E_m \), where the subset \( E_j=\{(O_{ij}, O_{ij+1}) | i=1, 2, \ldots, m-1 \} \cup (O_{im}, 1) | i=1, 2, \ldots, n \}. \) It expresses the non-repetition edges which connect the operations processed on the same machine \( j \). If the number of the operations processed on machine \( j \) is \( d_j \), the number of edges on subset \( E_j \) is \( \frac{n(n-1)}{2} \).

If we can fix the directions of non-directed edges on each subset and form a path over all nodes on each subset, then we can determine the sequence of the operations processed on each machine. These non-directed edges are replaced with directed edges (indicated with directed dashed line in Fig. 2) after the directions are fixed. Thus, we have found a schedule to the corresponding JSSP. Therefore, selecting a schedule is equivalent to choosing a graph \( G(N,AE) \), where \( E=E_1 \cup E_2 \cup \ldots \cup E_m \), \( E_j \) is a directed connected path which traverse all vertices associated with \( E_j \). The number of the edges on the subset \( E_j \) is \( d_j \). If the selected graph is acyclic, it corresponds to a feasible schedule.

Each vertex \( O_i \in N \) is corresponding to a weight \( t_i \), where \( t_i \) is the processing time of the \( j \)th operation of the \( j \)th job. Vertices 0 and 1 have weight zero. We define the longest path from vertex 0 to vertex 1 as the critical path. Then the length of critical path corresponds to the maximum completion time of the scheduling. Finding the minimum critical path from all the directed graph \( G(N,AE) \) is equivalent to finding the optimal solution of the JSSP (i.e. the solution corresponding to the minimum makespan).

![Fig. 1. The model of the problem.](image-url)
Consider a 3-job and 3-machine problem given in Table 1 as an example. A graph \( G=\langle N,A,E \rangle \) can be corresponding to this problem. From the above discussion, we get

\[
N = \{0,1,O_{11},O_{12},O_{13},O_{21},O_{22},O_{23},O_{31},O_{32},O_{33}\}
\]

\[
A = \{(0,O_{11}),(0,O_{12}),(0,O_{13}),(0,O_{21}),(0,O_{22}),(0,O_{23}),(0,O_{31}),(0,O_{32}),(0,O_{33})\}
\]

\[
E = E_1 \cup E_2 \cup E_3
\]

where \( E_1 = \{(O_{11},O_{21}),(O_{11},O_{31}),(O_{12},O_{22}),(O_{12},O_{32}),(O_{13},O_{23}),(O_{13},O_{33})\}, E_2 = \{(O_{12},O_{31}),(O_{12},O_{32}),(O_{13},O_{32}),(O_{13},O_{33})\}, E_3 = \{(O_{22},O_{13}),(O_{22},O_{33})\} \). Then we find a schedule to this problem as shown in Fig. 2 (edges in \( E_i \) \( i=1,2,3 \) are indicated with directed dashed line).

### 3. A new hybrid genetic algorithm for JSSP

#### 3.1. Encoding and decoding

Defining an appropriate representation of individual is the most important and critical issue for constructing an efficient GA. As mentioned above, determining a schedule is equivalent to determine a graph \( G=\langle N,A,E \rangle \), where \( E = E_1 \cup E_2 \cup \cdots \cup E_n \), and \( E_i \) is a directed connected path which traverses all vertices associated \( E_i \) that is, if we can make sure the sequence of the operations processed on each machine, we can determine a schedule. Thus, the following encoding scheme proposed by Falkenauer and Bouffoix [26] is adopted in this paper: a chromosome is composed of \( m \) substrings, corresponding to \( m \) different machines. Each substring represents the sequence of the operations processed on each machine, denoted by a permutation of integers from 1 to \( n \). The total length of a chromosome is \( n \times m \). Consider a 3-job and 3-machine problem given in Table 1 as an example. Suppose a schedule is given as shown in Fig. 2, then the corresponding chromosome is: \([2 \ 1 \ 3 \ 3 \ 1 \ 2 \ (2 \ 1 \ 3)]\). Here, substrings \([2 \ 1 \ 3], \quad (3 \ 1 \ 2), \quad (2 \ 1 \ 3)\) represent, respectively, the sequences of the operations processed on machine 1, 2 and 3. Conversely, from the chromosome \([2 \ 1 \ 3 \ (3 \ 1 \ 2) \ (2 \ 1 \ 3)]\), we also can get the corresponding schedule as shown in Fig. 2.

If the selected graph \( G=\langle N,A,E \rangle \) is acyclic, it corresponds to a feasible schedule. We can use the following method to decode all the feasible schedules. We introduce the decoding method through decoding the chromosome \([2 \ 1 \ 3 \ (3 \ 1 \ 2) \ (2 \ 1 \ 3)]\). According to Table 1, operations \( O_{11},O_{12},O_{13} \) must be scheduled on machine 1, operations \( O_{21},O_{22},O_{23} \) must be scheduled on machine 2 and operations \( O_{31},O_{32},O_{33} \) must be scheduled on machine 3, where \( O_j \) represents the \( j \)th operation of the \( i \)th job. And the schedule must meet the process constraint \( O_{11} \rightarrow O_{12} \rightarrow O_{13}, \quad O_{21} \rightarrow O_{22} \rightarrow O_{23}, \quad O_{31} \rightarrow O_{32} \rightarrow O_{33} \), where \( O_{ij} \rightarrow O_{ip} \) means operation \( O_j \) must be scheduled before \( O_{ip} \). The substrings \([2 \ 1 \ 3], \quad (3 \ 1 \ 2), \quad (2 \ 1 \ 3)\) represent the sequences of the operations processed on machine 1, 2 and 3, respectively. Therefore, the sequences of the operations processed on machine 1 is \( O_{21} \rightarrow O_{11} \rightarrow O_{12}, \quad O_{22} \rightarrow O_{21} \rightarrow O_{23}, \quad O_{31} \rightarrow O_{32} \rightarrow O_{33} \). So the first operations which can operate on the machine 1, 2, 3 are operations \( O_{21},O_{11},O_{22} \), respectively. After considering the process constraint, we schedule operation \( O_{21} \) on machine 1, operation \( O_{11} \) on machine 2 and operation \( O_{22} \) on machine 3 one after another at the earliest allowable time. Then the operations which can operate on machine 1, 2, 3 are, respectively, operations \( O_{21},O_{12},O_{31} \). After considering the process constraint, we schedule operations \( O_{11},O_{12},O_{13} \) on machine 1, 2, 3 one after another at the earliest allowable time. And then the operations can operate on machine 1, 2, 3 are, respectively, operations \( O_{21},O_{22},O_{31} \). After considering the process constraint, we schedule operation \( O_{12} \) on machine 1, operation \( O_{22} \) on machine 2 and operation \( O_{13} \) on machine 3 one after another at the earliest allowable time. Then we get the corresponding schedule shown in Fig. 3.

Obviously, if the selected graph \( G=\langle N,A,E \rangle \) includes one or more cycles, it corresponds to an infeasible schedule. Then we use the following method to change an infeasible schedule into a feasible schedule. Suppose a schedule is given as shown in Fig. 4, then the corresponding chromosome is: \([2 \ 3 \ 1 \ (1 \ 3 \ 2) \ (2 \ 1 \ 3)]\). So the
sequences of the operations processed on machine 1 is $O_{21} \rightarrow O_{31} \rightarrow O_{11}$, machine 2 is $O_{12} \rightarrow O_{32} \rightarrow O_{22}$ and machine 3 is $O_{23} \rightarrow O_{13} \rightarrow O_{33}$ and the process constraint is same to above example. So the first operations which can operate on machine 1, 2, 3 are, respectively, operations $O_{11}, O_{12}, O_{13}$. After considering the process constraint, we schedule operation $O_{21}$ on machine 1, operation $O_{22}$ on machine 3 at the earliest allowable time, respectively. Then the operations which can operate on machine 1, 2, 3 are, respectively, operations $O_{12}, O_{13}, O_{14}$. After considering the process constraint, we find none of the operations meets the process constraint because operations $O_{11}, O_{12}$ and $O_{13}$ have not been scheduled yet. That means the schedule corresponding to this chromosome is an infeasible schedule. Then we consider the next operations which can operate on each machine. They are operations $O_{11}, O_{12}, O_{13}, O_{14}$, respectively, processed on machine 1, 2, 3, 4. After considering the process constraint, we schedule operation $O_{11}$ on machine 1, operation $O_{12}$ on machine 2, operation $O_{13}$ on machine 3 at the earliest allowable time, respectively. Then we schedule operation $O_{12}$ on machine 1, operation $O_{13}$ on machine 2 at the earliest allowable time. After that we schedule operation $O_{13}$ on machine 2, operation $O_{14}$ on machine 3 at the earliest allowable time. At last we schedule operation $O_{14}$ on machine 3. Then we get the corresponding schedule as shown in Fig. 3. Finally, we find the chromosome corresponding to the Fig. 3 is $[(2 \ 1 \ 3) (3 \ 1 \ 2) (2 \ 1 \ 3)]$. That is, we change the infeasible schedule $[(2 \ 3 \ 1) (1 \ 3 \ 2) (2 \ 1 \ 3)]$ into a feasible schedule $[(2 \ 1 \ 3) (3 \ 1 \ 2) (2 \ 1 \ 3)]$ through above method.

### 3.2. Fitness function and selection operation

Fitness function is defined so as to determine the quality of individuals and it is usually relevant to the objective function to be optimized. The greater the fitness of a chromosome is, the greater its survival probability. In this study, the fitness function is defined as $f(x) = 1/makespan$, where $x$ is any individual in the population.

The selection phase is responsible to choose individuals for reproduction. Most of the selection strategies of GA adopt the method that the probability of selecting an individual is proportional to its fitness value. This can easily lead to the "premature" convergence because of the population being filled with more similar individuals [27]. To avoid this problem, a mixed selection operator based on the fitness value and the concentration value was proposed in this paper. This operator increases the diversity of the population and can prevent the "premature" convergence to some extent. Given any two individuals $x = [(x_1, x_2, \ldots, x_n), \{x_{n+1}, x_{n+2}, \ldots, x_{2n}\}]$ and $y = [(y_1, y_2, \ldots, y_n), \{y_{n+1}, y_{n+2}, \ldots, y_{2n}\}]$, we first give the definitions of the similarly degree, the similarity and the concentration.

**Definition 1.** For job $i$ in the $p$th substrings (i.e., the $p$th parenthesis) of $x$ and $y$ ($i \in \{1, 2, \ldots, n\}$), there may be some jobs which are in a position after $i$ or before $i$ in both corresponding strings of $x$ and $y$. We call the number of these jobs as the similarity degree of job $i$ in the $p$th substrings on $x$ and $y$, and denoted as $SDE(i_p)$.

**Definition 2.** We called the sum of the similarity degrees of all jobs in substring $p$ ($p \in \{1, 2, \ldots, m\}$) as the similarity degree of substring $p$ on $x$ and $y$, and denoted as $SDS(p)$.

$$SDS(p) = \sum_{i=1}^{m} SDE(i_p).$$

The larger the similarity degree, the more similar the corresponding substrings in two individuals.

**Definition 3.** Suppose $SDE(i_p)$ is the similarity degree of job $i$ in the $p$th substrings on individuals $x$ and $y$. Then the similarity of $x$ and $y$ is defined as

$$s(x, y) = \frac{\sum_{p=1}^{m} \sum_{i=1}^{n} SDE(i_p)}{n \times (n-1) \times m}.$$ 

We use the 4-job and 3-machine problem as an example to explain the similarity degree. Assume that $x = [(1 \ 2 \ 4 \ 3) (3 \ 1 \ 4 \ 2)]$ and $y = [(2 \ 1 \ 4 \ 3) (1 \ 4 \ 3 \ 2)]$ are two given individuals. By comparison of the corresponding substrings of these two individuals, we observe that: for substring 1 of both individuals, there are two jobs 3 and 4 in a position after 1. There are also two jobs 3 and 4 in a position after 2. There are three jobs 1, 2 and 4 in a position before 3, and there are two jobs 1 and 2 in a position before 4 and there is one job 3 in a position after 4. Therefore, $SDE(1, 1) = 2$, $SDE(2, 1) = 2$, $SDE(3, 1) = 3$, $SDE(4, 1) = 3$ and $SDS(1) = 2 + 2 + 3 + 3 = 10$. As similar, we get $SDS(2) = 3 + 3 + 3 + 3 = 12$ and $SDS(3) = 2 + 3 + 2 + 3 = 10$, respectively. Finally, the similarity of these two individuals is $s(x, y) = (10 + 12 + 10)/36 = 0.889$.

**Definition 4.** Suppose $L$ is the population size and $\lambda$ is a threshold value. In the current population, we use $Q(x, \lambda)$ to denote the total number of the individuals whose similarity with $x$ is greater than or equal to the threshold $\lambda$. Then the concentration value of individual $x$ is defined as $c(x) = Q(x, \lambda)/L$.

Obviously, the greater the value $s(x, y) \in [0, 1]$ is, the more similar the individuals $x$ and $y$. If $s(x, y) = 1$, the individuals $x$ and $y$ are the same individual. The greater the value $c(x) \in [0, 1]$ is, the more individuals are similar with individual $x$.

Suppose $x_1, x_2, \ldots, x_l$ are the individuals in the population, $L$ is the population size, $f(x_l) (l = 1, 2, \ldots, L)$ is the fitness function and $c(x_l)$ is the concentration value of individual $x_l$. The following selection operator based on the fitness value and the concentration value was given in this paper.

**Algorithm 1.** (A mixed selection operator based on the fitness value and the concentration value): 

**Step 1:** For each individual $x_i$ in the population, compute the pre-selection probability by using two different ways: one is by $p_i(x_i) = f(x_i) / \sum_{i=1}^{L} f(x_i)$, and another is by $p_i(x_i) = (1 - c(x_i)) / (L - \sum_{i=1}^{L-1} c(x_i))$. The former is based on fitness value and the later is based on the concentration. If $L = \sum_{i=1}^{L} c(x_i)$, we define $p_i(x_i) = 0$.

**Step 2:** For each individual $x_i$ in the population, define the selection probability $p_i(x_i) = \mu p_i(x_i) + (1 - \mu) p_i(x_i)$ called mixed selection probability by mixing pre-selection probabilities $p_i(x_i)$ and $p_i(x_i)$, where $\mu \in (0, 1)$ is a coefficient.

**Step 3:** Use the roulette wheel selection strategy with the probability $p_i(x_i)$ to select individuals from the population and generate new population.

### 3.3. A new crossover operator

The goal of the crossover is to obtain better offspring by making use of the parental information. In this paper, a new crossover operator based on the characteristic of the JSSP itself was designed. Suppose $x$ and $y$ are two parent individuals, the crossover operator designed in this paper can be described as follows.

**Algorithm 2.** (A new crossover operator based on the machine):

**Step 1:** Divide the machines 1, 2, \ldots, $m$ into two complementary sets $A$ and $B$. 
Step 2: For two parent individuals \(x\) and \(y\), swap the genes in the substrings that belong to set \(B\) with probability of \(p_c\) and get two children \(x'\) and \(y'\).

Swapping the genes in the substrings that belong to set \(B\) is equivalent to swapping the sequences of the operations processed on machines which belong to set \(B\). In this way, the children can effectively inherits the sequences of the operations processed on machines that belong to set \(A\).

We used a 3-job and 3-machine problem as an example to explain the steps of the crossover operator.

Suppose the two parent individuals parent 1 and parent 2 are 

\[
[(213) (312) (213)] \text{ and } [(321) (321) (321)].
\]

We get the graph model of the two parents shown in Figs. 5 and 6 at first.

Second, we divide the machines into two complementary sets, e.g., \([1, 3]\) and \([2]\). Third, we swap the substring 2 and get two children 

\[
[(213) (321) (213)] \text{ and } [(321) (312) (321)].
\]

Figs. 7 and 8 show the child 1 and 2 respectively.

The crossover operator designed in this paper is similar to the crossover operator in [26]. However, the crossover operator in [26] is like to a single point crossover, it can only separate the machines \(1, 2, \ldots, m\) from the middle to two set, such as \([1, 2, \ldots, m_1]; [m_1 + 1, m_1 + 2, \ldots, m]\) and then executes the crossover. The crossover operator designed in this paper is like to a multi-point crossover, it can randomly divide the machines \(1, 2, \ldots, m\) into two complementary sets. Obviously, the crossover operator designed in this paper is more flexible, and it increases the diversity of the population.

From the proposed crossover operator, we can observe that the children can inherit the sequences of the operations processed on some machines of their parents and the crossover operator is easy to execute.

3.4. A new mutation operator

The mutation operator further carries out some adjustments for a parent individual and gets an offspring. It can increase the diversity of the population and is helpful to jump out local optimal solutions.

According to Dell’Amica and Tubian [28], there are the following properties for JSSP:

1. If \(x\) is a feasible solution, then reversing one of the oriented edges on a critical path of \(x\) can never lead to an infeasible solution;
2. If the reversal of an oriented edge of a feasible solution \(x\) that does not belong to a critical path leads to a feasible solution \(x'\), then the critical path in \(x\) cannot be shorter than the critical path in \(x\).

According to these properties, a mutation operator based on the critical path was designed. Let \(PM(O_i)\) and \(SM(O_i)\) denote the immediate predecessor and the immediate successor of operation \(O_i\), respectively, on the same machine if they exist, otherwise, \(PM(O_i) = 0, SM(O_i) = 1\), where \(O_i\) represents the jth operation of the ith job. Then the proposed mutation operator can be described as follows.

Algorithm 3. (A new mutation operator based on the critical path):

Step 1. Calculate the critical path of a parent. Suppose that the obtained critical path is \(CP = \{O_{i_1}, O_{i_2}, \ldots, O_{i_k}\}\). Let \(k = 1\), the child = the parent.

Step 2. If operations \(O_{i_k}\) and \(O_{i_{k+1}}\) are processed on the same machine, then consider all possible permutations of \((PM(O_{i_k}), PM(O_{i_{k+1}}))\) and \((SM(O_{i_k}), SM(O_{i_{k+1}}))\) in which \(arc(O_{i_k}, O_{i_{k+1}})\) is inverted, randomly select one of them, and use this operations sequence to replace the previous sequence in the child with probability of \(p_m\).

Step 3. If \(k < q\) then let \(k = k + 1\) and go to step2, else, output the child.

In order to calculate the critical path, we presented a new algorithm for finding the critical path. Suppose \(S(O_i)\) and \(C(O_i)\) are the starting time and the completing time of operation \(O_i\), where \(O_i\) represents the jth operation of the ith job. \(i = 1, 2, \ldots, n, j = 1, 2, \ldots, m\). The algorithm is as follows.

Algorithm 4. (A new algorithm for finding the critical path):

Step 1: Decode the given chromosome to get a feasible schedule.

Step 2: Let \(K = \{O_i | i = 1, 2, \ldots, n; j = 1, 2, \ldots, m\}\), set \(Q = \emptyset\) (\(\emptyset\) represents an empty set).

Step 3: While \(\exists x \in K\), there is no \(\beta \in K\) fulfilling \(C(\alpha) = S(\beta)\), then delete \(\alpha\) from \(K\) and update \(K\).
3.5. A new local search operator

To enhance the speed of the proposed algorithm, a new local search operator is proposed in this paper. Assume that \( x \) is a given individual and \( x' \) is the resulting individual through local search operator from \( x \), and \( \sigma \) is the number of the new solutions to be generated in our local search operator.

Algorithm 5. (A new local search operator):

**Step 1:** Let \( k = 1 \), \( N_k = \emptyset \) (\( \emptyset \) represents an empty set).

**Step 2:** If \( k \leq \sigma \), go to step 3, else, go to step 6.

**Step 3:** Randomly select a machine number \( m_k \in \{1, 2, \ldots, m\} \).

**Step 4:** Change the sequence of the operations processed on machine \( m_k \) and gets a new solution \( N_k \).

**Step 5:** Put solution \( N_k \) into set \( N_k, k = k + 1 \), and then turn to step 2.

**Step 6:** Choose the best individual from \( N_k \) as \( x' \).

3.6. The framework of the HGA

Based on aforementioned genetic operators, we proposed a new hybrid genetic algorithm (briefly HGA) to solve the JSSP. The framework of the algorithm can be described as follows:

Algorithm 6. (Hybrid genetic algorithm: HGA):

**Step 1.** Initialization: Generate initial population \( P(0) \), and let \( t = 0 \). Evaluate the fitness value of each individual in \( P(0) \).

**Step 2.** Crossover: Group all the individuals in \( P(t) \) into pair wise randomly, and use Algorithm 2 to each pair of individuals and obtain a set of offspring, denoted by \( P'(t) \).

**Step 3.** Mutation: Use Algorithm 3 to all the individuals in \( P'(t) \) and obtain a set of offspring, denoted by \( P''(t) \).

**Step 4.** Evaluate fitness value: Evaluate the fitness value of each individual in \( P(t) \).

**Step 5.** Local Search: Use Algorithm 5 to each individual in \( P(t) \) with probability of \( p_l \) and obtain a set of offspring, denoted by \( P''(t) \).

**Step 6.** Selection: Select \( N \) individuals from \( P''(t) \) by Algorithm 1 to get a tentative population, still denoted as \( P''(t) \), then use the elitist strategy to the union of \( P'(t) \) and \( P(t) \) to get the next generation population \( P(t + 1) \).

**Step 7.** If the stop criterion is satisfied, then stop. Otherwise, let \( t = t + 1 \), and turn to step 2.

3.7. Convergence analysis of the HGA

At first, we introduce several symbols.

- \( \Omega \): The feasible solution space.
- \( x \): Any feasible solution, \( x \in \Omega \).
- \( f(x) \): The objective function, \( f(x) = \text{makespan} \).
- \( f^* \): The globally minimum of the objective function, i.e., \( f^* = \min \{ f(x) | x \in \Omega \} \).
- \( X^* \): The set of optimal solutions, i.e., \( X^* = \{ x \in \Omega | f(x) = f^* \} \).
- \( t \): The first time (generation) to find the optimal solution, i.e., \( T = \min \{ t | f^*_t = f^* \} \), where \( f^*_t \) is the minimum of the nth generation of population.

For proposed algorithm, we can get the following conclusions.

Theorem 1. For algorithm HGA with any initial population \( P(0) \), suppose the crossover, mutation, local search and selection operators are independent of each other and the crossover probability is \( p_c \), the mutation probability is \( p_m \) and the local search probability is \( p_l \). HGA can find the global optimum in limited iterations with probability 1, i.e., \( p(T < \infty) = 1 \).

Proof. According to Dell’Amico and Trubian [28], for \( \forall x \in \Omega \), it is possible to construct a finite sequence of mutation operations, which will lead from \( x \) to a globally minimal solution \( x^* \in \mathcal{X} \). This process can be represented as \( x = x_0 \rightarrow x_1 \rightarrow \cdots \rightarrow x_{t-1} \rightarrow x^* \). Suppose the critical path of \( x_l \) is \( c_p = (O_{i_0}, O_{i_1}, \ldots, O_{i_{l-1}}) \). Form Algorithm 3, if operations \( O_{i_0}, O_{i_1}, \ldots, O_{i_{l-1}} \) are processed on the same machine, then consider all possible permutations of \((P(M(O_{i_0}, O_{i_1}, \ldots, O_{i_{l-1}})), O_{i_0}, O_{i_1}, \ldots, O_{i_{l-1}}))\) and \((O_{i_p}, O_{i_q}, \ldots, SM(O_{i_p}, O_{i_q}, \ldots))\) (a total of six cases) in which arc \((O_{i_p}, O_{i_q}, \ldots, SM(O_{i_p}, O_{i_q}, \ldots))\) is inverted, randomly select one of them, and use this operations sequence to replace the previous sequence with probability of \( p_m \).

Therefore, the probability of getting \( x_k \) from \( x_{l-1} = (k = 0, 1, 2, \ldots, t-1) \) after mutation should fulfill \( p(x_k = \text{mutation}(x_{l-1})) \geq (p_m(1 - p_m)^{t-1}) = \delta_m > 0 \), where \( q \) is the number of operations on the critical path. Note \( k_s \) as the length of the shortest path from \( x \) to a point in \( X^* \) through a finite successive mutation. Let \( k^* = \max_k \{ k_s | x \in X^* \} \).
For all \( t \), the probability of \( x \) not being destroyed after crossover operator is \( p(x \in P(t)) \geq 1 - p_r - \delta_t > 0 \).

Thus, after mutation, the probability of getting \( x_1 \) from \( x \) is \( p(x_1 \in P(t)) \geq \delta_m > 0 \). The probability of \( x_1 \) not being destroyed after local search operator is \( p(x_1 \in P(t)) \geq 1 - p_l - \delta_t > 0 \). The probability of \( x_1 \) being selected to the next generation population \( P(t+1) \) is \( p(x_1 \in P(t+1)) \geq p_1(x_1) > 0 \). Let \( \delta = \min(p_1(x_1),p_2(x_2),...p_{n}(x_n)) \), then, after the crossover, mutation, local search and selection once, the probability of getting \( x_1 \) from \( x \) is \( p(x_1 \in P(t+1)) \geq \delta, \delta_m, \delta_t > 0 \).

Let \( p(\delta) \) denote the probability of getting \( x^* \in X^* \) from \( x \) after \( k \) generations, then \( p(\delta) \geq (\delta, \delta_m, \delta_t)^k > 0 \). Let \( p(k^*) \) denote the probability of getting an optimal solution from \( x \) after \( k \) generations, then \( p(k^*) \geq (\delta, \delta_m, \delta_t)^{k^*} > 0 \). The probability not finding the optimal solution after \( k^* \) generation is \( p_{not}(k^*) \leq 1 - \delta > 0 \). So \( p_{min}(1) \leq (1 - \delta)^{k^*} > 0 \) and \( p_{min}(k) \leq (1 - \delta)^{k^*k} > 0 \).

Therefore, \( \lim_{k \to \infty} p_{not}(k) \leq \lim_{k \to \infty} (1 - \delta)^{k^*k} = 0 \). Since \( p(T < + \infty) \geq (\text{find the optimal solution at } k \text{ generations}) = 1 - p_{not}(k) \), let \( k \to \infty \), we can get \( p(T < + \infty) \geq 1 \). Thus, \( p(T < + \infty) = 1 \). □

**Theorem 2.** Algorithm HGA converges to the optimal solution with probability 1.

**Proof.** According to Theorem 1, HGA can find the global optimum in limited iterations with probability 1. Because of adopting the elitist strategy in HGA, the algorithm converges to the optimal solution with probability 1. □

From the process proof, it can be easily found that because of using the new mutation operator based on the critical path designed in this paper, algorithm HGA can find the global optimum in limited iterations with probability 1. Therefore, if we do not use the mutation operator designed in this paper, the convergence of the algorithm cannot be guaranteed.

### 4. Simulation results

In order to verify the good performance of the proposed hybrid genetic algorithm, we use 43 instances from two classes of standard JSSP test problems: Fischer and Thompson [2] instances FT06, FT10, FT20 and Lawrence [29] instances LA01–LA40.
The HGA was compared with some algorithms reported in literature [16,30–35] in recent years.

The parameters used in simulations are as follows: the population size $N=100$, the crossover probability $p_c=0.7$, the mutation probability $p_m=0.1$, and the local search probability $p_l=0.5$. In the selection operator, $\lambda = 0.8$ and $\mu = 0.7$. In the local search operator, $\sigma = 10$. The algorithm was run 10 independent times on each test problem.

The HGA was implemented in Visual C++ on a computer with a 2.66 GHz Intel Core 2 Duo CPU. Table 2 shows the experimental results. It lists problem name, problem size (number of jobs x number of operations), the best known solution (BKS), and the solution obtained by each of the compared algorithms. Among them, column HGA1 lists the results obtained by using the local search operator in the HGA and column HGA2 lists the results obtained by not using the local search operator.

Table 2 shows that the proposed algorithm is able to find the best known solution for 33 instances, i.e., in about 77% of the instances. For small problems FT06, FT10, FT20 and LA01–LA15, almost all the algorithms can find the optimal solution. For relatively large problems LA16–LA40, the results of the proposed algorithm (HGA) are better than or equal to those of other algorithms. For HGA, the results obtained by using the local search operator are better than the results obtained by not using the local search operator. This shows that the proposed local search operator takes an important role in improving the quality of the solutions.

For each algorithm, we can use formula $RD = 100 \times (\text{MFM} - \text{BKS}) / \text{BKS}$ for each instance to calculate the relative deviation, where MFM means the minimum makespan found and BKS means the best known solution. We use ARD to denote the average value of relative deviations for all the instances. Table 3 shows the number of instances solved (NIS), and the average relative deviation (ARD). The ARD was calculated for the Hybrid Genetic Algorithm (HGA), and for the other algorithms (OA) listed in the table. The last column (Improvement) presents the reduction of ARD obtained by the HGA with respect to each of the other algorithms.

From Table 3, we know the deviation of the minimum makespan found from the best known solution of HGA is only on average 0.17%. The proposed algorithm yields a significant improvement in solution quality with respect to all other algorithms.

5. Conclusions

To solve the JSSP more effectively, a mixed selection operator based on the fitness value and the concentration was designed in order to increase the diversity of the population. Crossover operator based on the machine, and mutation operator based on the critical path were designed according to the graph model of JSSP. To calculate the critical path, a new algorithm was presented. A local search operator was designed in order to improve the quality of the solutions. Based on these, a hybrid genetic algorithm was proposed, and the convergence of HGA to the global optimal solution in finite generation with probability 1 is proved. Then the convergence of HGA to the optimal solution with probability 1 is proved. The experimental results show that the proposed algorithm is effective and performs better than the compared algorithms.

Acknowledgment

The work is supported by National Natural Science Foundation of China (60873099) and Ph.D. Programs Foundation of Ministry of Education of China (20090203110005).

Table 3

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>NIS</th>
<th>ARD</th>
<th>Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OA%</td>
<td>HGA%</td>
<td>HGA%</td>
</tr>
<tr>
<td>Yang (2008) MA</td>
<td>13</td>
<td>0.06</td>
<td>0.04</td>
</tr>
<tr>
<td>Goncalves (2005) Param. active</td>
<td>43</td>
<td>0.39</td>
<td>0.17</td>
</tr>
<tr>
<td>Ombuki (2004) LSGA</td>
<td>25</td>
<td>5.22</td>
<td>0.29</td>
</tr>
<tr>
<td>Cielo (2003) AIS</td>
<td>24</td>
<td>1.51</td>
<td>0.15</td>
</tr>
<tr>
<td>Wang (2001) MGA</td>
<td>11</td>
<td>0.27</td>
<td>0.13</td>
</tr>
<tr>
<td>Binato (2001) GRASP</td>
<td>43</td>
<td>1.68</td>
<td>0.17</td>
</tr>
<tr>
<td>Sabuncuoglu (1999) Beam search</td>
<td>41</td>
<td>4.02</td>
<td>0.18</td>
</tr>
<tr>
<td>Dauzere (1997) Tabu search</td>
<td>43</td>
<td>0.86</td>
<td>0.17</td>
</tr>
</tbody>
</table>

References


