Using novel particle swarm optimization scheme to solve resource-constrained scheduling problem in PSPLIB

Ruey-Maw Chen\textsuperscript{a,}\textsuperscript{*,} Chung-Lun Wu\textsuperscript{b}, Chuin-Mu Wang\textsuperscript{a}, Shih-Tang Lo\textsuperscript{c}

\textsuperscript{a}Department of Computer Science and Information Engineering, National Chin-yi University of Technology, Taichung 411, Taiwan, ROC
\textsuperscript{b}Department of Electrics Engineering, National Chin-yi University of Technology, Taichung 411, Taiwan, ROC
\textsuperscript{c}Department of Information Management, Kun Shan University, Tainan 710, Taiwan, ROC

\begin{abstract}
This investigation proposes an improved particle swarm optimization (PSO) approach to solve the resource-constrained scheduling problem. Two proposed rules named delay local search rule and bidirectional scheduling rule for PSO to solve scheduling problem are proposed and evaluated. These two suggested rules applied in proposed PSO facilitate finding global minimum (minimum makespan). The delay local search enables some activities delayed and altering the decided start processing time, and being capable of escaping from local minimum. The bidirectional scheduling rule which combines forward and backward scheduling to expand the searching area in the solution space for obtaining potential optimal solution. Moreover, to speed up the production of feasible solution, a critical path is adopted in this study. The critical path method is used to generate heuristic value in scheduling process. The simulation results reveal that the proposed approach in this investigation is novel and efficient for resource-constrained class scheduling problem.
\end{abstract}

\section{1. Introduction}
Scheduling notion has been widely applied in many fields, such as production planning (Zhai, Tiong, Bjorndsson, & Chua, 2006), process scheduling in operating systems (Shaw et al., 1999), classroom arrangement (Mathaisel & Comm, 1991), aircrew-scheduling (Chang, 2002), nurse scheduling (Ohki, Morimoto, & Miyake, 2006), food industrial (Simeonov & Simeonovova, 2002), control system (Fleming & Fonseca, 1991) and grid computing (Buyya, Abramson, & Giddy, 2000). Generally, these problems commonly accompany the cost considerations related to certain constraints; therefore, scheduling scheme plays an important role in obtaining solutions for constraints satisfaction and cost minimization.

The resource-constrained scheduling problem (Merkle, Middendorf, & Schmeck, 2002) is an optimization problem considering how to schedule the activities of a multiprocessor system such that the makespan of the schedule is minimized while satisfying given precedence constraints between the activities and resource requirements of the scheduled activities per time unit do not exceed the given different types of resources capacity limit. However, the minimum makespan is hard to be obtained since the inestimable situation of constraints. Most scheduling problems are confirmed to be NP-complete combinatorial problems, especially for large scale scheduling problems. There are some algorithms were studied to solve scheduling problems and obtain optimal solution. Branch-and-bound method (Jalilvand, Khammamadi, & Shabaninia, 2005) is able to find optimal solutions of scheduling problem. However, the execution time required is impractical when the number of activities increases. Comparatively, several greedy-based algorithms such as shortest job first (SJF) (Lupetti & Zagorodnov, 2006) and priority scheduling (Li, Bettatti, & Zhao, 1997) are able to solve the resource-constrained scheduling problem in feasible time, but it is hard to accommodate the algorithm to the changed problem’s constraints situation. Hence, the optimal solution is seldom obtained by greedy-based algorithms.

Hopfield and Tank (1985) first proposed a type of artificial neural network; named Hopfield neural network (HNN) for solving optimization problems. The HNN mimics the cooperation of all neurons in the brain (cerebrum) and learns to solve problems. We have applied HNN to scheduling multiprocessor job with resource and timing constraints (Huang & Chen, 1999). Furthermore, Chen, Lo, and Huang (2007) combined competitive scheme with slack neurons to solve real-time job scheduling problems. In HNN, the neurons’ output is decided based on the input information: synaptic weights and bias inputs from outside of the neuron. An energy (cost or fitness) function is adopted to confine the problems’ constraints and targets. The solution of a scheduling problem is yielded as the energy function converges. However, the energy function is not easy to be designed once the constraints of the scheduling problems become complicated.
Many studies solve the resource-constrained scheduling problem using the meta-heuristics based methods, such as genetic algorithm (GA) (Liu & Wang, 2005), simulated annealing algorithm (SA) (Bouleimen & Lecocq, 2003), tabu search (TS) (Thomas & Salhi, 1998), ant colony optimization (ACO) (Merkle et al., 2002), particle swarm optimization (Luo, Wang, Tang, & Tu, 2006), and others. The GA mimics the mechanism of natural selection as a global evolution (Holland, 1987); then a more superior solution part will be inherited via crossover operation, and increasing the diversity of solution via mutation process. Both crossover and mutation steps are used to expect obtaining a better solution of the studied problem.

Originally, simulated annealing was investigated by Kirkpatrick et al. as a stochastic method for combinatorial optimization problem (Kirkpatrick, Gelatt, & Vecchi, 1983). The optimal solution is a stable state when the thermal energy of the system minimized. Therefore, the required procedure is to decrease the thermal energy of the system by cooling down temperature parameter. The temperature lowered as the iteration goes. Restated, cooling the system to lower energy state. Moreover, the SA applies a probability to decide whether allows the system toward the higher energy state to avoid trapped on the local optimum. Tabu search is an approach to prevent the search from sinking into the local minimum. The key skill used in tabu search is to record the solutions which have been obtained. Therefore, the following search can avoid searching the same obtained solutions (Glover Fred., 1989).

The ACO emulates the foraging behavior of ants (Dorigo & Gambardella, 1997). The ant left pheromone on the trail of the searched path from nest to the destination. The pheromone deposited on the way is for other ants to identify and communicate with each other. Additionally, the amount of pheromone is inverse proportional to the length of path; a large amount of pheromone is accumulated at the shorter path. Accordingly, ants are toward the path with plenty of pheromone. The maximum amount of pheromone on the path can be regarded as an ant notification signal indicating where the shorter path is located at. Hence, the amount of pheromone increases right as more ants pass through the shortest path. It is a self-reinforce situation of searching the shortest path, this situation will result fast convergence and trap in local optimal solution. Therefore, an evaporation mechanism is applied in ACO to reduce the pheromone to obtain global optimal solution. Using ACO to solve multiprocessor system scheduling with precedence and resources constraints was proposed (Chen, Lo, Wang, & Wu, 2006; Chen, Zhang, Hao, & Dai, 2006).

The particle swarm optimization (PSO) is first proposed by Kennedy and Eberhart (1995). In PSO, a swarm of particles spreads in the space and the position of a particle represents a solution of a dedicated problem. Each particle would move to a new position based on the global experience of the swarm and the individual experience of the particle for the global optimum. The PSO has been applied to solve the scheduling problems. Luo et al. (2006) used PSO to solve resource-constrained project scheduling problem (PCPSP); they showed that the PSO is applicable to various combinatorial problems and scheduling problems. Zhang, Sun, Zhu, and Yang (2008) solved flowshop scheduling problem (FSP) in terms of the PSO, and Chen, Lo et al. (2006) and Chen, Zhang et al. (2006) solved task scheduling in grid based on PSO. Hence, this investigation further modifies the PSO by involving extra mechanisms to enhance the efficiency in solving the scheduling problem.

The characteristics of above stated algorithms are concluded as follows.

- Most of them require a fitness function to evaluate the quality and feasibility of the obtained solution.
- The features or characteristics of a good solution are referred to when constructing a new solution.
- In GA, a fragment of the good solution is combined with another fragment of another feasible solution to build a new solution.
- ACO and PSO algorithms infer new solution from the helpful experience, and are possible to produce premature convergence. Accordingly, these two algorithms will converge to the local optimal solution. Therefore, most suggested algorithms usually include additional techniques to escape from the premature convergence. For example, randomly disturbing the system or depressing the self-reinforce is applied (Selman, Kautz, & Cohen, 1994). Eventually, there are two aspects to improve the algorithm for scheduling global optimal solution; one is to enhance the heuristic ability (exploitation ability), the other is to strengthen the ability of escaping local optimum or expanding the searching area in the solution space (exploration ability).

This study improved PSO to solve scheduling problem for optima or near-optima schedule by minimizing makespan. This improved PSO uses the critical path method to enhance the heuristic ability (exploitation ability). The critical path method considers only logical dependencies between activities but resources availability. However, the critical path method still helps speeding up PSO to gain the feasible solution when the conflict of resources is not acute. Meanwhile, a delay local search to lead the solution is applied in this improved PSO to escape from the local optima (exploration ability). With delay local search, the scheduling system leaves some activities not to be activated occasionally and hence remains some unused resources available for other activities. Moreover, bidirectional scheduling combines forward and backward scheduling to expand the searching area in the solution space is integrated into this improved PSO (Li & Willis, 1992). The resources allocation for forward and backward scheduling is different, and hence two different schedules may be obtained, namely, schedules become diversified. The experimental simulations demonstrate that the approach of this study for the cases of RCPSP is able to obtain more optima or near-optima solutions.

This article is organized as follows. Section 2 introduces the RCPSP and PSO, and describes how to use PSO to solve RCPSP. Section 3 presents that the proposed novel PSO uses the critical path method to enhance the heuristic ability (exploitation), escapes from the local optima using delay local search (exploration) and expands the searching area in the solution space in terms of integrating the bidirectional scheduling. The simulated cases and results of experiments are displayed in Section 4. Finally, Section 5 presents the conclusions and discussions.

2. The resource-constrained scheduling problem and particle swarm optimization (PSO)

2.1. Resource-constrained scheduling problem

The scheduling application has been applied in various fields. Among them, the resource-constrained scheduling application is a general scheduling problem involving activities to be scheduled to some identical processors with precedence and resource constraints satisfied. Activities are confined to the various constraints and achieve a certain object. The studied scheduling problem with resource constraint in this investigation is defined as follows:

- The object of the scheduling is to find the minimal makespan schedule.
- There’re N activities in the multiprocessor system. Every activity requires a processing duration $d_j$ ($j = 1,...,N$). Meanwhile, activities are non-preemptive in the schedule.
• Activities have precedence constraints; a feasible schedule is an assignment of activities to processors such that an activity comes in the schedule when all its predecessors have finished. Let \( P_j (S_j) \) be the set of direct predecessors (relative to the direct successors) of activity \( j \). Activity 0 is the start activity that has no predecessor while activity \((N + 1)\) is the finishing activity that has no successor. The start activity and the finishing activity have no resource requirements for any type of resources and have no processing duration.

• There are different types of resources available in the system. Each type of resource has a limited quantity. Let \( Q \) be a set of \( q \) resource types and \( R_k (k = 1, \ldots, q) \) is the resource capacity for resource type \( k \). Every activity \( j \) requires various resources \( r_{j1}, r_{j2}, \ldots, r_{jq} \), where \( r_{jk} \) denotes the quantity of resource type \( k \) demanded by activity \( j \) when activity \( j \) is scheduled. Moreover, resource can be reallocated to other activities when that resource is released from finished activities.

• Moreover, the resource constraints also assumed that the total amount of a resource type \( k \) required by scheduled activities can not exceed \( R_k \) at any time, such that \( \sum_{j \in P_j (S_j)} r_{jk} \leq R_k \), where the \( S_j(t) \) is the set of activities scheduled by particle \( i \) at time \( t \). Therefore, the solution \( S_i \) contains different activities scheduled at different time.

2.2. The particle swarm optimization (PSO)

The particle swarm optimization (PSO) is a multi-agent general meta-heuristic method, and can be applied extensively in solving many tough problems. The PSO consists of a swarm of particles in the space; the position of a particle is indicated by a vector which presents a solution. PSO is initialized with a population of \( M \) particles randomly spreads in the space and then starts to search for the best position (namely, schedule herein). At each step or iteration, the local best of each particle and global best of the swarm are determined through evaluating the performances, i.e., the fitness values or makespan, of current distribution of particles. A particle moves to a new position as a new solution which is guided by the velocity (a vector). Therefore, the velocity plays an important role in creating a new solution.

There are two experience positions are used in the PSO. One is the global experience position of all particles of the swarm, which memorizes the global best solution obtained from all positions (solutions) of all particles. The other is the individual experience position of each particle, which memorizes the local best solution acquired from the positions (solutions) which have been passing through by the corresponding particle. These two experience positions combining the weighted previous velocities are used to determine the impact or influence on the current velocity. Re-stated, the current velocity retains part of previous velocity (called the inertia) and driving particle toward the direction based on the global experience position and the individual experience position. Accordingly, the particles can reach new positions (solutions) based on their own inertia and experiences.

Let an \( N \) dimension space (the number of dimension is corresponding to the components of solution) has \( M \) particles. For the \( i \)th particle \((i = 1, \ldots, M)\), the position (solution) consists of \( N \) components, i.e., \( X_i = \{X_{i1}, \ldots, X_{iN}\} \), where \( X_{ij} \) is the \( j \)th component of the position. The velocity of particle \( i \) is \( V_i = \{V_{i1}, \ldots, V_{iN}\} \), where \( V_{ij} \) is the \( j \)th component of the velocity. The individual experience is a position, \( L_i = \{L_{i1}, \ldots, L_{IN}\} \) denotes the local best solution for the \( i \)th particle. Moreover, \( G = \{G_1, \ldots, G_N\} \) represents the global best experience shared among all the population of particles achieved so far. The mentioned parameters above are used to update the \( j \)th component of the position and the \( j \)th component of velocity for the \( i \)th particle, as shown in Eq. (1).

\[
\begin{align*}
V_i^{new} &= w \times V_i + c_1 \times r_1 \times (L_i - X_i) + c_2 \times r_2 \times (G_j - X_i), \\
X_i^{new} &= X_i + V_i^{new}
\end{align*}
\]

where \( w \) is an inertia weight used to determine the influence of the previous velocity on the new velocity. The \( c_1 \) and \( c_2 \) are learning factors. They are used to drive the \( i \)th particle how to approach the new position either for close to the individual experience position or expecting close to global experience position. Furthermore, the \( r_1 \) and \( r_2 \) are the random numbers uniformly distributed in \([0, 1]\), controlling the tradeoff between the global and local exploration abilities during search. The values of \( w \), \( c_1 \) and \( c_2 \) are assumed to be positive herein.

3. Improved particle swarm optimization

The procedure of the proposed improved particle swarm optimization is shown as in Table 1. The delay local search rule and bidirectional scheduling rule will be discussed in Section 4.

For particle \( i \), each activity \( j \) has a defined priority \( pr_{ij} > 0 \) \((j = 1, \ldots, N)\) for deciding which activity will be selected. Generally, the activity with the higher priority will be selected first, and assign to one processor. In this proposed PSO, the activity’s priority \( pr_{ij} \) determination is based on the \( X_i^{new} \).

3.1. Heuristic function design with critical path consideration

To speed up convergence of the algorithm for problem, the critical path is involved in this study. The critical path is computed without the consideration of resources constraints. One activity has a higher heuristic value, while the time is close to over the activity’s latest starting time obtained from critical path. Re-stated, the critical path is applied to determine the heuristic value which hence affects the probability based on greedy property for selecting activity \( j \) of particle \( i \) at time \( t \), and is used in the activity selection rule. The heuristic value definition is given in Eq. (2).

The critical path method in this study is used to obtain the activity’s earliest starting time \( Est_{stj} \), latest starting time \( Lst_{stj} \) and the difference \( D_j \) between \( Lst_{stj} \) and \( Est_{stj} \). The \( D_j \) is defined as the heuristic value for scheduling activity \( j \) of \( i \)th particle at time \( t \). Intrinsically, the \( D_j \) is defined for the timing constraints for activity \( j \), to meet the timing requirement. The activity \( j \) is unable to be selected at time \( t \) when the time \( t \) is earlier than activity’s earliest starting time, then the \( D_j \) is set to zero.

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pseudo code of proposed novel PSO (NPSO).</td>
</tr>
</tbody>
</table>

1. Initialize
2. Loop
3. \( i = 1 \)
4. The bidirectional scheduling rule decides using forward scheduling or backward scheduling solution
5. Loop
6. \( i = i + 1 \)
7. Initialize \( t = 0 \) and \( j(t = 0) \)
8. Loop
9. Loop
10. Select one activity \( f \in j(t) \) with activity selection rule
11. If delay local search is activated then either continue or delay activities
12. else schedule activity \( f \) start at time \( t \)
13. \( j(t) = j(t) \cup \{f\} \)
14. Until \( f = \emptyset \)
15. \( t = t + 1 \)
16. Until \( j(t) = \emptyset \)
17. Update the local or global optimal solution if necessary
18. Until all particles have built a complete solution
19. Updating position and velocity of every particle by Eq. (1)
20. Until End condition is reached
Furthermore, the activity \( j \)'s earliest starting time is ahead of current time \( t \) but the latest starting time is behind the time \( t \), the \( g_{ijt} \) is inverse proportional to the \( D_i \). The \( g_{ijt} \) increases as the \( D_i \) increases when the time \( t \) is already later than the activity \( j \)'s latest starting time. Restated, the activity \( j \) is associated with higher \( g_{ijt} \) if the time \( t \) approaches the latest starting time of activity \( j \).

Moreover, the notion of state transition rule of ACO is utilized to generate the new solution in this study. The state transition rule used in ACO is computed by the pheromone ad heuristic value. Accordingly, the heuristic value applied in this proposed novel PSO would affect the probability of selecting activity \( j \) of particle \( i \) at time \( t \).

\[
C_{ijt} = \begin{cases} 
0, & \text{if } Est_{st_j} > t \\
\frac{1}{\left(\prod_{j \in J_i(t)} D_j\right)} \cdot \frac{1}{g_{ijt}}, & \text{if } Est_{st_j} \leq t < Lstst_j, \ j \in J_i(t) \\
\frac{1}{\prod_{j \in J_i(t)} D_j}, & \text{otherwise} 
\end{cases} 
\]

where \( D_j = \left| Lstst_j - t \right| \) and \( c \) is a large enough constant.

3.2. Priority decision of activity

In this proposed PSO, the new solution generation is based on activity priority. The priority \( p_{rj} \) is calculated based on the inverse of the solution \( X_{ijt}^{new} \). The \( X_{ijt}^{new} \) is the start processing times of activity \( j \) for new solution of particle \( i \). The new solution (position) generated by particle is not used directly for schedule, but transformed to a priority. Therefore, a smaller \( X_{ijt}^{new} \) has a higher priority \( p_{rj} \) indicating the new solution suggests the activity \( j \) with high priority needs to be started earlier. Restated, a larger \( p_{rj} \) indicates that the activity \( j \) is possible selected earlier. The priority of activity is defined as listed in Eq. (3).

\[
p_{rj}^{new} = \left( X_{ijt}^{new} \right)^{-1} 
\]

3.3. Activity selection rule and resource consideration

The priority \( p_{rj} \) is extended to define \( prcg_{ijt} \) under resource constraints and other consideration. For the resource constraints, each activity \( j \) is associated with a constraint state value \( c_{ijt} (j = 1, \ldots, N) \) for particle \( i \) at time \( t \). The value of \( c_{ijt} \) is decided to indicate the resources availability of remaining resources at time \( t \), since some resources are already allocated to scheduled activities. Restated, the \( c_{ijt} \) is a constraint state value based on the resource constraints at time \( t \). This constraint state value is set to zero \((c_{ijt} = 0)\), once activity \( j \) is prohibited to be selected at time \( t \) due to that the required resources are more than remaining resources. Otherwise, the constraint state value is set to one, \( c_{ijt} = 1 \). Restated, activity \( j \) at time \( t \) is a candidate activity when \( c_{ijt} = 1 \). Meanwhile, \( J_i(t) \) is a set consisting of all candidate activities which can be assigned to run at time \( t \). An activity is included in \( J_i(t) \) based on the precedence among activities. An activity is then excluded from \( J_i(t) \), if the activity is selected for execution.

The probability, \( Pb_{ijt} \), of activity \( j \in J_i(t) \) used for activity selection at time \( t \) is determined by \( prcg_{ijt} \) (the product of \( p_{rj} \), \( c_{ijt} \), and \( g_{ijt} \)). The \( Pb_{ijt} \) is a normalized \( prcg_{ijt} \) as shown in Eq. (4). Moreover, the activity selection rule at time \( t \) is designed as shown in Eq. (5).

\[
Pr_{rcg_{ijt}} = p_{rj} \times c_{ijt} \times g_{ijt} \\
Pb_{ijt} = \begin{cases} 
\frac{\prod_{j \in J_i(t)} prcg_{ijt}}{\sum_{j \in J_i(t)} prcg_{ijt}}, & \text{if } j \in J_i(t) \\
0, & \text{otherwise} 
\end{cases} \\
\]

\[
j' = \begin{cases} 
\arg \max_{j \in J_i(t)}\{prcg_{ijt}\}, & \text{if } q < q' \text{ and } \sum_{j \in J_i(t)} c_{ijt} \neq 0 \\
S, & \text{if } q = q' \text{ and } \sum_{j \in J_i(t)} c_{ijt} \neq 0 
\end{cases} 
\]

The activity \( j' \) is selected from \( J_i(t) \) according to the product of \( pr_{rj} \) and \( g_{ijt} \) values, for those activity \( j \) with \( c_{ijt} = 1 \), then the scheduling process is processed until \( J_i(t) \) is empty. The activity selection is decided by maximum \( prcg_{ijt} \) when \( q < q' \). Otherwise activity \( j \) is chosen according to \( S \) value. \( S \) is a random variable selected according to the probability distribution, \( Pb_{ijt} \), as given in Eq. (4). The \( q \) is a random number uniformly distributed in \([0, 1]\), and the \( q' \) is a predefined parameter in \([0, 1]\). The \( g_{ijt} \) value will be higher, if the time \( t \) is closing to the latest starting time of activity \( j \), or the time \( t \) is over the latest starting time based on Eq. (2), then a higher \( prcg_{ijt} \) will be induced. Accordingly, the activity \( j \) will have the higher probability \( Pb_{ijt} \) to be selected.

3.4. Solution representation and schedule generation

The position vector of each particle presents a solution; all the components of the position vector are the start times of all activities. If there are \( N \) activities in the scheduling system, the dimension of the space in which the particles move is \( N \). Thus, the \( N \) components of the position vector present \( N \) start times of \( N \) activities. The schedule generation procedure for particle \( i \) of this investigation is summarized as Fig. 1.

In the new solution generation rule of this proposed PSO, the start time of activities \( X_{ijt} \) and priority \( (p_{rj}) \) of activities are updated according to Eqs. (1) and (3) respectively. Afterward, the particles construct new schedule using activity selection rule Eq. (5). Furthermore, the individual (or global) experience is replaced by the new solution when the new solution is better than existing individual (or global) experience.

4. Solution searching mechanisms

In this work, two solution space searching mechanisms are applied. One is called delay local search rule which is used to alter some activities in schedule while generating new solution. The other is the bidirectional scheduling rule which employs alternate forward with backward scheduling to expand the searching area in the solution space.

\[
\begin{align*}
V_{ijt}^{new} &= w \times V_{ijt} + c_1 \times r(L_{ijt} - X_{ijt}) + c_2 \times r(G_{ijt} - X_{ijt}) \\
X_{ijt}^{new} &= X_{ijt} + V_{ijt}^{new}, \ j = 1, \ldots, N \\
p_{rj}^{new} &= \left(X_{ijt}^{new}\right)^{-1} \\
prcg_{ijt} &= p_{rj} \times c_{ijt} \times g_{ijt} \\
Pb_{ijt} &= \frac{\prod_{j \in J_i(t)} prcg_{ijt}}{\sum_{j \in J_i(t)} prcg_{ijt}}, \ \text{if } j \in J_i(t) \\
0, & \text{otherwise} \\
\end{align*}
\]

Select \( j' \) by activity selection rule, \( j' = 1, \ldots, N \)

Generate schedule (Deciding the start processing time of job \( j' \))
4.1. Delay local search rule

Generally, the allocation of resources in the scheduling process is to have as many activities executed as possible at any time, and the resulting schedule is sound. However, this way may cause some resources occupied by activities for a span of time, and lead to some crucial activities delaying start at an inappropriate time for optimal scheduling (minimum makespan). Restated, this delay is mainly caused by the shortage of certain resources occupied by previous scheduled activities. For instance, if one activity requiring many resources at time \( t \) is selected for execution, then some other activities requiring the same resources will be prohibited from being selected to process for a period of time. Accordingly, these delayed activities result in a larger makespan than that of the optimal solution. Hence, this study involves a mechanism of delay local search to delay some activities being selected from \( Ji(t) \). This mechanism would leave remaining resources available and these available resources can be allocated to other activities properly. Restated, one activity can be processed later to let the other activities be processed ahead to yield the global optimal solution under the resource constraints. This delay local search is indicated in the step 11 of the proposed novel PSO in Table 1.

The delay local search rule as illustrated in Eq. (6). Using this delay local search, some resources requested by selected activity will be reserved for other activities. An activity is delayed if \( q < q_0 \), where \( q \) is the random number uniformly distributed in \([0, 1] \), and \( q_0 \) is a predefined small value \((0 < q_0 < 1)\). The \( q_0 \) is defined as a delay rate to indicate the delay probability. Restated, an activity \( j \) selected from \( Ji(t) \) will not be assigned to processor at time \( t \) once the activity is being selected and \( q < q_0 \). This mechanism is similar to the mutation operation of genetic algorithm (GA).

\[
delay(j, t) = \begin{cases} 
  \text{true} & \text{if } q / C < q_0 \\
  \text{false} & \text{otherwise}
\end{cases}
\]

Fig. 2 depicts an example case for 13 activities with precedence and resource constraints. Figs. 3–6 all use the forward scheduling mechanism. Figs. 3 and 4 are the schedule results with and without delay local search applied, the resources allocation is also shown in Fig. 5. Fig. 6 lists the remaining available resources for each type of resources for each time slot with and without using the delay local search.

Suppose \( S_i(t) \) is the set of activity scheduled by particle \( i \) at time \( t \). In the example of Fig. 2 and the resulting schedule in Fig. 3, \( S_1(1) = \{2, 3, 4\} \) at time 1, the activity 2 will be finished at time 3. Activity 10 would be assigned for processing at time 4 according to the precedence, i.e., \( S_2(4) = \{3, 4, 10\} \). The candidate activities at time 5 are represented by the set \( Ji(5) = \{3, 6, 7, 10\} \). Nevertheless, the set of activity scheduled by particle \( i \) at time 5 is \( S_2(5) = \{3, 6, 7\} \) since the total amount of \( R_3 \) resources available is 12, which is insufficient for processing activities 6 and 10 concurrently. And, activity 6 and 10 must be mutually excluded at the same time. Hence, the activity 6 has to be delayed to start until activity 10 finish at time 13.

However, if the delay local search is applied appropriately at time 4, the activity 10 is excluded from being assigned for running. Accordingly, the set of the selected activities at time 5, \( S_2(5) = \{3, 6, 7\} \) is possible, and the activity 6 can start earlier. Restated, activities 6 and 10 are not allowed to use resource \( R_3 \) simultaneously, based on the resource requirement. Moreover, activities 5 and 6 need more resource \( R_4 \) than system provided, hence, activities 5
and 6 are prohibited to execute at the same time. Consequently, the application of delay local search increases the resource utilization and hence is able to yield shorter makespan solution. If the delay local search is activated at time 4 and put away the selection of activity 10, the makespan is 34 as shown in Fig. 4, otherwise, the makespan is larger than 34 as displayed in Fig. 3.
4.2. Bidirectional scheduling rule

The general scheduling scheme allots activities to processors to yield solution while satisfying all the required constraints such as resources, precedence and other constraints. Generally, the system has different resources allocation would cause diverse solutions. Hence, Li and Willis (1992) suggested applying bidirectional scheduling to discover other solutions, since the situations of resources requested by activities in forward and backward scheduling are different, the resulting schedules are different. The forward scheduling and backward scheduling would search for solutions in different searching area of the solution space.

Some scheduling cases are appropriate by forward scheduling, and some scheduling cases are suitable using backward scheduling. Hence, this study alternates forward scheduling with backward scheduling during iteration to increase the scheduling efficiency. In this study, two particle swarms are employed; one consists of forward scheduling particles and the other comprises of backward scheduling particles. Then, the forward scheduling particle works on the ordinary problem instance, and the backward one works on the reversed problem instance. Restated, the directed paths of the precedence graph are reversed. Forward and backward particles work separately on their own position and velocity matrices. Although the solutions generated by backward scheduling are not always better than the one by forward scheduling. However, the backward scheduling provides an opportunity to diversify the solutions and increase the opportunity to obtain the optimal solution.

The simulation results of the example case (illustrated in Fig. 2) scheduled by bidirectional scheduling without delay local search are displayed in Figs. 7–9. The Gantt charts of resulting schedule only by using backward scheduling (Makespan = 34, 100 iterations, 10 particles, 3 processors: P1, P2, P3 are used). The makespan of the resulting schedule using backward scheduling is 34, and the makespan is larger than 34 for the forward scheduling as shown in Fig. 3. Restated, the optimal schedule can be obtained by backward scheduling.

5. Experimental examples and results

This study simulates the Single Mode Data Sets cases in PSPLIB [29] library, these cases consist of 30–120 activities. The simulation parameters used in this investigation are set as listed in Table 2. The feasible solutions (the position of particles) are maintained by every activity’s start processing time during iteration. The objective of the studied algorithm is to find the schedule with minimal makespan. The simulation program is coded by C language and running at the PC with Pentium4 3.4 GHz CPU.

The following simulations were applied to verify how many cases of PSPLIB library can be solved to yield optimal solution or lower bound solution by the proposed novel PSO. In PSPLIB library, the benchmark optimal or lower bound solution for each instance is also included. There is no processor constraint is given in PSPLIB. Hence, in the simulations, the number of processor is assumed sufficient for all instance cases.

There are 5 approaches are compared in this study and as listed in Table 3, they are proposed novel particle swarm optimization for scheduling (denoted by PSO+), PSO with bidirectional scheduling (indicated by PSO Bidirectional), PSO with delay local search (PSO Delay), traditional PSO (PSO), and ant colony optimization for scheduling as in (Chen et al., 2006; Chen et al., 2006) is denoted by ACO.

Figs. 10–13 demonstrate the simulation results of all 480 instances for 30, 60, 90 activities cases and 600 instances for 120 activities case. To evaluate the studied algorithm, each instance was run for 10, 100, and 1000 iterations. These figures show that the proposed novel PSO (denoted by PSO+, combining the critical path method, delay local search and bidirectional scheduling method) is able to find more optimal schedules than the other classes of PSO and ACO algorithms.

Table 4 illustrates the 30 activities example instance j301_6 in PSPLIB. Fig. 14 display the case j301_6 simulation results and total distributions of the resources used by the PSO and PSO+ methods.
In this instance case, there are 4 types of resource, 30 activities and sufficient processors available are assumed. The makespan of scheduling result using PSO is 49 and the makespan of scheduling result by PSO+ is 48 (optimal). Fig. 14a shows the total resource usage comparison by using PSO and PSO+ methods. Fig. 14b and c display the resulting schedules using traditional PSO and PSO+ respectively. Moreover, the simulation results of the 120 activities instance (j12022_4 in PSPLIB) using the PSO and PSO+ methods are demonstrated in Fig. 15. The simulation results of case j12022_4

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of particles: ( M )</td>
<td>10</td>
</tr>
<tr>
<td>Initial positions of particles: ( X_i )</td>
<td>Random value</td>
</tr>
<tr>
<td>Initial priority: ( p_{ri} )</td>
<td>Random value</td>
</tr>
<tr>
<td>Initial velocity of particles: ( V_i )</td>
<td>0</td>
</tr>
<tr>
<td>Inertia weight: ( w )</td>
<td>0.7</td>
</tr>
<tr>
<td>Learning factors: ( c_1 ) and ( c_2 )</td>
<td>0.7</td>
</tr>
<tr>
<td>Predefined parameter for delay rate: ( q_0 )</td>
<td>0.05</td>
</tr>
<tr>
<td>Parameter in heuristic value determination: ( c )</td>
<td>50</td>
</tr>
<tr>
<td>Parameter in activity selection: ( q' )</td>
<td>0.95</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Critical path method</th>
<th>Delay local search</th>
<th>Bidirectional scheduling</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSO+</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>PSO bidirectional</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>PSO delay</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>PSO</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>ACO</td>
<td>○</td>
<td>×</td>
</tr>
</tbody>
</table>

Fig. 9. The resource allocation of each type of resources.
are similar to that of case j301_6. Restated, the PSO+ method can yield optimal makespan of the solution schedule.

The scheduling problems in PSPLIB have 30, 60, 90, and 120 activities cases. The 30, 60 and 90 activities problem cases each has 480 instances, 120 activities problem case has 600 instances. To further evaluate the studied algorithm, each instance was run for 1000 iterations; and each problem size case was tested 10 trials.

The assessment of different approaches is represented by the solution quality. The solution quality is commonly measured by the relative percentage deviation (RPD):

\[
RPD = \frac{\sum_{\text{instances}} (100\% \times \frac{\text{Obtained} - \text{Best}}{\text{Best}})}{\text{instances}}
\]  

(7)

\[
ARPD = \frac{\sum_{\text{trials}} RPD}{\text{trials}}
\]

(8)

\[
BRPD = \min\{RPD\}
\]

(9)

In Eq. (7), the Obtained is the makespan of solution obtained by certain approach for certain instance, best is the optimum solution or the lower bound of this instance in the library. Therefore, RPD of
Table 5

<table>
<thead>
<tr>
<th>Problem size/approach</th>
<th>ARPD (%) PSO+</th>
<th>ARPD (%) PSO bidirectional</th>
<th>ARPD (%) PSO delay</th>
<th>ARPD (%) PSO</th>
<th>ARPD (%) ACO</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>0.54</td>
<td>0.84</td>
<td>1.03</td>
<td>2.07</td>
<td>1.57</td>
</tr>
<tr>
<td>60</td>
<td>3.71</td>
<td>3.80</td>
<td>4.04</td>
<td>5.26</td>
<td>4.34</td>
</tr>
<tr>
<td>90</td>
<td>3.79</td>
<td>3.89</td>
<td>4.42</td>
<td>5.61</td>
<td>4.32</td>
</tr>
<tr>
<td>120</td>
<td>9.38</td>
<td>9.38</td>
<td>11.39</td>
<td>14.00</td>
<td>10.21</td>
</tr>
<tr>
<td>Average</td>
<td>4.36</td>
<td>4.48</td>
<td>5.22</td>
<td>6.73</td>
<td>5.11</td>
</tr>
</tbody>
</table>
certain problem size is calculated by averaging the difference between Obtained and best of all instances. Since each problem size case was simulated 10 trials, an average relative percentage deviation (ARPD) is defined as shown in Eq. (8) and a best relative percentage deviation (BRPD) is defined as listed in Eq. (9).

Table 5 lists the comparison of ARPD among different approaches for different scale problems. For example, the ARPD, an average of 4.36% and a maximum of 9.38% were achieved using proposed PSO+. Meanwhile, the comparison of BRPD among different approaches is given in Table 6.

According to Tables 5 and 6, the experiment results on different problem sizes by proposed PSO+ are better than that by other approaches.

Moreover, to verify the effect on the solution of using the proposed delay local search, different delay rates were investigated in this work. The delay rate is similar to the mutation rate in GA method. The ARPD and BRPD of the simulation results using various delay rates are shown in Tables 7 and 8.

6. Conclusions and discussion

This study proposes a novel PSO scheme to solve precedence and resource-constrained scheduling problems. A critical path method is applied to facilitate the new activity selection in Eqs. (4) and (5). Moreover, a delay local search mechanism (Eq. (6)) is utilized to escape from local optimum solution. Furthermore, this investigation applied bidirectional scheduling mechanism of alternate forward and backward scheduling to expand the searching area in the solution space. The simulation results reveal that this improved novel PSO is effective and efficient to solve the class of PSPLIB scheduling problems. The number of optimal solutions yielded by proposed novel PSO is obvious superior to other traditional methods such as ACS and PSO as shown in Figs. 10–13. The more optimal solutions found as iteration increased. Especially, for the problem size of 30 activities instances, the proposed method is able to find 450 optimal solutions in 480 instances (93.75%) when 10,000 iterations tested. Meanwhile, the PSO+ approach derived form PSO outperforms ACO approach in PSPLIB problems. It can be seen that the PSO+ integrated with appropriate search methods in this study can effectively improve the algorithm’s efficiency as listed in Tables 5 and 6.

Moreover, a small value of the delay rate is suggested to find sound solution quality. Essentially, delay local search is to stir a small disturbance in solution space. However, a small fixed delay rate used in large scale problem would cause large disturbance since large scale problem has large amount of solution population. Hence, to yield small BRPD, the larger scale problem requires the smaller delay rate as proved in Tables 7 and 8.

Furthermore, some other important features of this study are summarized below.

1. The variation of resources usage is an important factor in the scheduling algorithms. Therefore, make efficient use of resources shorten the resulting schedule makespan.
2. The suggested delay local search makes some determined activities delayed, and is able to have the better resource utilization of the system. Hence, the order of processing activities is changed and therefore yields the resulting optimal solution.
3. Additionally, bidirectional scheduling alternatively applying forward and backward scheduling mechanism can further increase resource utilization efficiency while searching for the optimal solution in different area in the solution space. Therefore, optimal solution can be yielded.
4. One of the important characteristics about the scheduling algorithms is its efficiency, the computation complexity. The execution time of this improved PSO is proportional to the number of particles M and activities N. For each particle, a solution requires N executions of activity selection rule, and activity selection rule would calculate the probability of selection for each activity (also N executions). Restated, the computation complexity of O(M × N²) for each iteration is provided.

This study mainly studies the scheduling problems with precedence and resource constraints, and the processors are assumed to be identical. More complex conditions or situations should be further considered such as the setup time between activities of a certain machine, or if there is communication cost between two activities which are processing in different processors. Moreover, in a dynamic situation, there may be some emergency activities arriving at a certain time or changing the resources available and requirement. Meanwhile, the design of heuristic value is also possible to be improved for better solution.

References


