A Self-adaptive Iterated Local Search Algorithm on the Permutation Flow Shop Scheduling Problem

Xingye Dong1, Maciek Nowak2, Ping Chen3 and Youfang Lin1
1Beijing Key Lab of Traffic Data Analysis and Mining, School of Computer and IT, Beijing Jiaotong University, Beijing 100044, China
2Quinlan School of Business, Loyola University, Chicago, IL 60611, U.S.A.
3TEDA College, NanKai University, Tianjin 300457, China

Keywords: Scheduling, Permutation Flow Shop, Total Flow Time, Iterated Local Search, Self-adaptive Perturbation.

Abstract: Iterated local search (ILS) is a simple, effective and efficient metaheuristic, displaying strong performance on the permutation flow shop scheduling problem minimizing total flow time. Its perturbation method plays an important role in practice. However, in ILS, current methodology does not use an evaluation of the search status to adjust the perturbation strength. In this work, a method is proposed that evaluates the neighborhoods around the local optimum and adjusts the perturbation strength according to this evaluation using a technique derived from simulated-annealing. Basically, if the neighboring solutions are considerably worse than the best solution found so far, indicating that it is hard to escape from the local optimum, then the perturbation strength is likely to increase. A self-adaptive ILS named SAILS is proposed by incorporating this perturbation strategy. Experimental results on benchmark instances show that the proposed perturbation strategy is effective and SAILS performs better than three state of the art algorithms.

1 INTRODUCTION

Since the pioneering work of Johnson (Johnson, 1954), the permutation flow shop problem (PFSP) has attracted considerable attention. In this problem, there are $n$ jobs and $m$ machines, and each job has $m$ operations. The jobs need to be processed on $m$ machines in the same sequence, that is to say no pre-emption is allowed. The $i$th operation of each job needs to be processed on the $i$th machine. All the jobs are available at time zero and each machine can serve at most one job at any time. Any operation can be processed only if its previous operation has been processed and the requested machine is available. The PFSP is NP-complete when minimizing total flow time with more than one machine (Garey et al., 1976).

For the purpose of finding high-quality solutions within a reasonable computation time, many methods, including simple heuristics and more complex metaheuristics (Dong et al., 2013), have been proposed. Among the existing metaheuristics, local search procedure plays an important role. Several ant colony algorithms are proposed by Rajendran et al. (Rajendran and Ziegler, 2004; Rajendran and Ziegler, 2005), working with well designed local search procedures and performing better than some constructive heuristics by Liu and Reeves (Liu and Reeves, 2001). Tasgetiren et al. (Tasgetiren et al., 2007) apply a particle swarm optimization (PSO) algorithm by using the smallest position value rule. They also propose a hybrid algorithm with variable neighborhood search (VNS), called PSO$_{VNS}$, and it performs quite well on Taillard’s benchmark instances (Taillard, 1993). Pan et al. (Pan et al., 2008) develop two metaheuristics, a differential evolution algorithm and an iterated greedy algorithm hybridized with a referenced local search procedure. Local search procedures are also used in several genetic algorithms (Tseng and Lin, 2009; Zhang et al., 2009; Tseng and Lin, 2010). Recently, Tasgetiren et al. (Tasgetiren et al., 2011) designed an artificial bee colony algorithm and a discrete differential evolution algorithm. Both algorithms use a local search procedure taking advantage of the iterated local search (ILS) by Dong et al. (Dong et al., 2009) and the iterated greedy (IG) algorithm by Ruiz and Stützle (Ruiz and Stützle, 2007). An asynchronous genetic local search algorithm, embedding an enhanced variable neighborhood search, is also addressed by Xu et al. (Xu et al., 2011) for the PFSP minimizing total...
flow time.

Local search procedure is used as an embedded procedure in the aforementioned metaheuristics. However, it is also used in an iterative way to form a metaheuristic called iterated local search (ILS). Dong et al. (Dong et al., 2009) develop an ILS to solve the PFSP minimizing total flow time, in which the perturbation method swaps several pairs of adjacent jobs and the perturbation strength, denoted by the number of swapping pairs, is evaluated. A multi-restart iterated local search (MRSILS) algorithm is proposed by Dong et al. (Dong et al., 2013), improving the perturbation method mainly by generating restart solutions from a set of elite solutions. Their experiments show that the MRSILS increases the performance of the methodology used in Dong et al. (Dong et al., 2009) significantly, while performing comparably to or better than five other state of the art metaheuristics (Pan et al., 2008; Zhang et al., 2009; Tasgetiren et al., 2011). Costa et al. (Costa et al., 2012b) study the insertion neighborhood. Six different combinations in total are calibrated. Later, they extend this work by developing an algorithm hybridizing VNS and path-relinking on a particle swarm framework, with promising experimental results on Taillard’s benchmark instances (Costa et al., 2012a). Pan and Ruiz (Pan and Ruiz, 2012) propose two local search methods based on the well known ILS and IG frameworks. They also extend them to population-based versions; however, their experiments show that the two local search methods perform better than the population-based versions.

Though ILS has performed well on the PFSP minimizing total flow time, one limitation is that the search process often cannot improve the best solution, even with dozens of iterations. For example, this occurs with the MRSILS by Dong et al. (Dong et al., 2013), although the pooling strategy in this algorithm augments the ability to find better solutions. A reason for the search process to be limited is that the perturbation method moves randomly selected jobs to other randomly selected positions without considering the search status. In order to improve perturbation quality, this work proposes a method to evaluate the convergence status of the search. With a greater focus on convergence, the probability increases that a job is moved to a position where a solution with larger objective can be generated. Based on this observation, a self-adaptive ILS (SAILS) is proposed and evaluated. Comparison results with the MRSILS by Dong et al. (Dong et al., 2013) and two local search based algorithms by Pan and Ruiz (Pan and Ruiz, 2012) on Taillard’s benchmark instances (Taillard, 1993) show that the new perturbation method leads to improved performance.

The remainder of this paper is organized as follows. In Section 2, the formulation of the PFSP with total flow time criterion is presented. Section 3 describes the evaluation method and illustrates the proposed algorithm. The evaluation method is analyzed and SAILS is compared with several state of the art algorithms in Section 4, then the paper is concluded in Section 5.

2 PROBLEM FORMULATION

In this paper, the PFSP is discussed with the objective of minimizing total flow time. This problem is an important and well-known combinatorial optimization problem. In the PFSP, a set of jobs \( J = \{1, 2, \ldots, n\} \) available at time zero must be processed on \( m \) machines, where \( n \geq 1 \) and \( m \geq 1 \). Each job has \( m \) operations, each of which has an uninterrupted processing time. The processing time of the \( k \)th operation of job \( j \) is denoted by \( p_{ij} \), where \( p_{ij} \geq 0 \). The \( k \)th operation of a job is processed on the \( j \)th machine. An operation of a job is processed only if the previous operation of the job is completed and the requested machine is available. Each machine processes these jobs in the same order and at most one operation of each job can be processed at a time. The PFSP to minimize total flow time is usually denoted by \( F_{m|prmu}\Sigma C_j \) (Pinedo, 2001), where \( F_{m} \) describes the environment, \( prmu \) is the set of constraints and \( C_j \) denotes the completion time of job \( j \). Let \( \pi \) denote a permutation on the set \( J \), representing a job processing order. Let \( \pi(k), k = 1, \ldots, n \), denote the \( k \)th job in \( \pi \), then the completion time of job \( \pi(k) \) on each machine \( i \) can be computed through a set of recursive equations:

\[
C_{i,\pi(1)} = \sum_{r=1}^{i} P_{r,\pi(1)} \quad i = 1, \ldots, m
\]

\[
C_{1,\pi(k)} = \sum_{r=1}^{k} P_{1,\pi(r)} \quad k = 1, \ldots, n
\]

\[
C_{i,\pi(k)} = \max\{C_{i-1,\pi(k)}, C_{i,\pi(k-1)}\} + p_{i,\pi(k)} \\
\quad i = 2, \ldots, m; k = 2, \ldots, n
\]

Then \( C_{\pi(k)} = C_{m,\pi(k)} \), \( k = 1, \ldots, n \). The total flow time is \( \sum_{k=1}^{n} C_{\pi(k)} \), or the sum of completion times on machine \( m \) for all jobs. The objective of the PFSP when minimizing total flow time is to minimize \( \sum C_{\pi(k)} \), or \( C_\pi \) for short.

3 THE PROPOSED ALGORITHM

According to Lourenço et al. (Lourenço et al., 2010), the general framework of ILS has four key compo-
nents: the method generating the initial solution, the local search procedure, the acceptance criterion and the method to perturb a solution. In this work, the H(2) heuristic by Liu and Reeves (Liu and Reeves, 2001) is used to generate the initial solution, as it can generate an initial solution in negligible time and has been used by Dong et al. (Dong et al., 2013; Dong et al., 2009) to form quite good ILS algorithms. As for the other three key components, the local search procedure and the evaluation method used to indicate the convergence status are discussed in Section 3.1. The acceptance criterion and the perturbation method are addressed together in Section 3.2. Finally, the proposed ILS algorithm is illustrated in Section 3.3.

### 3.1 Evaluation Method and Local Search

In this work, a solution to the discussed problem is presented as a permutation of the jobs and the commonly used insertion local search is chosen as the local search procedure. During the local search, each job is removed from its original position and inserted into the other $n-1$ positions to see whether the solution can be improved. Suppose the current local optimum solution is $\pi$ and $f_{\pi(j)}$ denotes the objective value of the solution generated by removing job $\pi(i)$ and inserting it into position $j$. Among the $n-1$ solutions, the best one is chosen and its objective is denoted by $f^*_{\pi(j)}$. The objective values form a matrix $F$, called the objective matrix. For all of the jobs, the average objective value of $f^*_{\pi(j)}$, $i = 1, \ldots, n$ can be computed as:

$$avg_X = \frac{1}{n} \sum_{i=1}^{n} f^*_{\pi(i)},$$

(4)

This average objective value can be used as an indicator of the convergence status. Suppose the current best objective value in the search process is $f_b$. The larger the difference $avg_X - f_b$, the worse the solutions in the neighborhood surrounding $\pi$ and the more difficult it is to escape from the local minimum in this neighborhood. This is an indicator that the perturbation strength should be increased. In order to temper the influence of this indicator as the difference between $avg_X$ and $f_b$ becomes larger, this difference is adjusted such that:

$$avg'_X = (avg_X - f_b)^{1/k},$$

(5)

where $k$ is an integer, and its value is tuned in Section 4.

The local search used in this work is shown in Figure 1, where $\pi$ denotes the start solution, $\pi^*$ denotes the best objective value in the search process is $\pi^*$.

#### 3.2 Acceptance and Perturbation Method

In the general framework of the ILS (Lourenço et al., 2010), a decision is made whether to accept a solution as the local optimum when it is reached, and then the solution may be perturbed to generate a restart solution that continues the local search procedure. Dong et al. (Dong et al., 2013) propose a pooling strategy that leads the search to a more promising solution space, generating highly competitive solutions. In this work, the pooling strategy is also used, while the perturbation method is adapted by using the indicator $avg'_X$ and incorporating the concept of temperature control from simulated annealing (Nikolaev and Jacobson, 2010).

According to Dong et al. (Dong et al., 2013), a set of elite solutions is pooled during the local search. Suppose the set of elite solutions is pool and $\pi$ is the solution chosen from it. It is perturbed by randomly selecting a job, $i$, and moving it to another randomly selected position with probability $\exp(-avg'_X/T)$, where $T$ denotes temperature. In this work, $T$ is a constant value. Alternatively, with probability $1 - \exp(-avg'_X/T)$ the randomly selected job is moved to a position that can generate a worse solution. The probability of moving job $i$ to position $j$, where $j \neq i$, is denoted by $p_j$ and can be determined
procedure Perturbation()
1. if \( |pool| < pool\_size \) then
2. \( \pi \leftarrow \pi^*; F_\pi \leftarrow F_\pi^* \);
else
3. Choose a solution from pool as \( \pi \), get
the corresponding \( F_\pi \) from \( F\_pool \);
endif
4. Compute \( \text{avg}_\pi \) using \( F_\pi \);
5. Randomly select a job \( \pi(i) \) from \( \pi \);
6. if \( \text{rand}(\cdot) < \exp(-\text{avg}_\pi/T) \) then
7. Move \( \pi(i) \) to another randomly selected
position in \( \pi \) to form a new \( \pi \);
else
8. Move \( \pi(i) \) to another position in \( \pi \) accor-
ding to Eq. (6) to form a new \( \pi \);
endif
9. return \( \pi \);
end

Figure 2: Pseudo code of the perturbation method.

Table 1: Global variables used in the perturbation method.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \pi^* )</td>
<td>The best solution found so far</td>
</tr>
<tr>
<td>( \pi )</td>
<td>A variable denotes a solution</td>
</tr>
<tr>
<td>pool</td>
<td>The set of elite solutions</td>
</tr>
<tr>
<td>pool_size</td>
<td>The size limitation of pool</td>
</tr>
<tr>
<td>( F_\pi )</td>
<td>Objective matrix corresponding to ( \pi )</td>
</tr>
<tr>
<td>( F_\pi^* )</td>
<td>Objective matrix corresponding to ( \pi^* )</td>
</tr>
</tbody>
</table>
| \( F\_pool \) | Pool of objective matrix correspond-
ing to pool |

using a roulette methodology, such that the probability increases with the cost of the solution generated
by the move. This probability is determined as:

\[
p_j = \frac{\sqrt{f_\pi(i,j) - f_b + 1}}{\sum_{k=0, k \neq j} \sqrt{f_\pi(i,k) - f_b + 1}} \tag{6}
\]

The perturbation method is illustrated in Figure 2.
In this figure, \( \text{rand}(\cdot) \) is a function that generates a random number uniformly distributed in the range \([0, 1]\), and the global variables are listed in Table 1.

3.3 Self-adaptive ILS

The proposed ILS is named self-adaptive ILS (SAILS), with the pseudo code presented in Figure 3.

In this work, the \( pool\_size \) is set to 5, as this value
is found to be effective in Dong et al. (Dong et al., 2013), although the parameter is rather robust. The
parameter \( k \) (used in Eq. (5)) and temperature \( T \) are
tuned in Section 4.1. There are usually two termi-
nation criteria in the literature: a maximum number of
iterations and limited computational time. In this

work, the latter is applied on line 5 in order to more
easily compare the SAILS with other algorithms.

In considering insertion local search, the number of
solutions to be evaluated is at least \( n \times (n - 1) \)
and the data complexity for evaluating one solution
is \( O(nm) \), so the time complexity of the local search
is at least \( O(n^2m) \). In literature, the CPU time
limitation is often set to \( \rho \times m \) milliseconds, where \( \rho \) is a
constant (Taggetiren et al., 2007; Ruiz and Stützle,
2007; Xu et al., 2011; Costa et al., 2012b; Costa
et al., 2012a; Pan and Ruiz, 2012). This setting has
a shortcoming in that smaller instances can run with
more CPU time relative to larger instances. For ex-
ample, an instance with 20 jobs and 20 machines
may have \( 1000 \times 20 \times (20 - 1) \) solutions evaluated
in \( \rho \times 20 \times 20 \) milliseconds CPU time, i.e. checking all
the jobs 1000 passes; while an instance with 500 jobs
and 20 machines still has the same number solutions
evaluated in \( \rho \times 500 \times 20 \) milliseconds CPU time, i.e.
checking all the jobs only about 1.52 passes. In or-
der to avoid this unfair, the CPU time limitation is set
to \( \rho n^2m \) milliseconds. In this work, \( \rho \) is set to 0.004,
0.012 and 0.02, respectively.

4 COMPUTATIONAL RESULTS

In this section, the proposed algorithms are evaluated.
Firstly, two parameters of the SAILS are tuned, then
the SAILS is compared with state of the art algorithms
and shown the effective of the newly proposed pertur-
bation method.

The benchmark instances used for analysis are
Table 2: Tuning the parameters of the SAILS.

<table>
<thead>
<tr>
<th>temperature $T$</th>
<th>value of $k$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>0.306</td>
</tr>
<tr>
<td>4</td>
<td>0.309</td>
</tr>
<tr>
<td>5</td>
<td>0.318</td>
</tr>
</tbody>
</table>

The performance of the algorithms are tested using the relative percentage deviation (RPD), which is calculated as:

$$RDP = \frac{(C - C_{\text{best}})}{C_{\text{best}}} \times 100$$

where $C$ is the result found by the algorithm being evaluated and $C_{\text{best}}$ is the best result provided by Pan and Ruiz (Pan and Ruiz, 2012), found as the best solution among their four proposed algorithms and 11 other state of the art works.

Each algorithm is implemented in C++, running on three similar PCs, each with an Intel Core2 Duo processor (2.99 GHz) and 2GB main memory. Though each computer has two processors, only one is used in the experiments, as no parallel programming technique has been applied.

4.1 Tuning of the SAILS

There are two parameters in the proposed SAILS, the first one is $k$, used in Eq. (5); the other is the temperature $T$, used in the self-adaptive perturbation (see Fig. 2). The $k$ is set to 1, 2, 3 and 4, respectively, and the $T$ is set to 3, 4 and 5, respectively. So there are 12 combinations in total. The terminal criterion is set to $0.02 \times n m$. As it is too time consuming for the 500 jobs instances, the SAILS is only run on the first 110 instances. The overall averaged RPDs (ARPD) for these 12 combinations are listed in Table 2. From this table, it can be seen that the results are quite similar for each pair of $T$ and $k$. However, the results are generally better with $k = 2$. As the performance is one of the best cases with $k = 2$ and $t = 4$, we choose them in the following experiments.

4.2 Evaluation of the SAILS

The values for the average RPD (ARPD) of all 120 instances are presented in Tables 3 - 5 for each combination of $n$ and $m$. From these tables, it can be seen that SAILS generally performs the best. And with the prolonging of CPU time, the superiority of the SAILS increases (See Fig. 4). The reason for this is that the self-adaptive strategy will be applied more times in the search with prolonged CPU time, and then resulting significantly better performance for the SAILS. This also shows that the proposed strategy is effective, particularly with longer CPU time.

This phenomenon can also be observed on large instances. With the 500 job instances, when the CPU time limitation is set to $0.004 \times n m$, SAILS performs the worst. When the CPU time limitation is set to $0.012 \times n m$, SAILS surpasses MRSILS and PR-IGA, but is still worse than PR-ILS. With the CPU time limitation set to $0.020 \times n m$, SAILS further outperforms MRSILS and PR-IGA, and is quite similar to PR-ILS.

Figure 4: Comparison results for the SAILS with the MRSILS, PR-ILS, PR-IGA and SAILS (0.004m^3 ms CPU time).
5 CONCLUSIONS

Iterated local search algorithms are powerful for solving the PFSP minimizing total flow time, with the perturbation method playing an important role. The MRSILS algorithm by Dong et al. (Dong et al., 2013) is a state of the art ILS algorithm. However, one shortcoming is that the best current solution cannot be improved in many local search runs during the search process. Further, the perturbation method only moves a randomly selected job to another randomly selected position, without any bias, such that it is difficult to escape from a “deep” local optimum.

In order to overcome this shortcoming, a self-adaptive perturbation method is proposed in this paper. In this method, the search status is evaluated by calculating the average objective value of a sample of the neighborhoods around the local optimum. The greater the difference between the average objective value and the best current objective value, the higher the probability that the perturbation strength is increased and a randomly selected job is moved to a position where worse solutions can be generated. The SAILS algorithm is proposed based on the above analysis. Experimental results on benchmark instances show that SAILS works quite well, especially for long CPU times. The proposed methodology developed here may potentially be applied to other problems, as escaping local minimums with local search is a difficulty for many combinatorial optimization problems.

ACKNOWLEDGEMENTS

This work is supported by The Fundamental Research Funds for the Central Universities of China (Project Ref. 2014JBM034, Beijing Jiaotong University).

REFERENCES


