OUTSOURCING DECISIONS IN SINGLE MACHINE SCHEDULING PROBLEM WITH MULTIPLE EXTERNAL FACILITIES

Jung Man Hong and Jong Hyup Lee

Key words: outsourcing, single machine scheduling, due date, pseudo-polynomial.

ABSTRACT

This paper considers outsourcing decisions in a scheduling problem. The objective of the problem is to minimize the outsourcing costs under the constraints of the capacity of outsourcing facilities and due date of each job. The problem is composed of three kinds of decisions. The first decision is the selection of jobs to be processed in-house, the second is to schedule the in-house processing jobs with due date constraints, and the last is to select an outsourcing provider for each outsourced job where the outsourcing provider has a capacity constraint. Some optimality conditions and solution properties for the problem are presented. A solution algorithm with pseudo-polynomial complexity is suggested to find the optimal solution of the problem. The main contributions of this paper are as follows: (1) the mathematical model for the problem is proposed, (2) the pseudo-polynomial algorithm, Depth First Search (DFS) is developed to find the optimal solution and (3) several optimality properties for the problem are addressed. Numerical experiments show that the DFS algorithm has better results than Dynamic Programming (DP) in efficiency even for larger numbers of jobs and outsourcing providers.

I. INTRODUCTION

The proposed problem is to find the optimal schedule with outsourcing in which jobs are processed in external facilities to improve the overall scheduling quality. The objective of the problem is to minimize the total outsourcing cost under the constraints of the capacity of external facilities and the due date of each job.

The scheduling problem with outsourcing would be considered in a project in which a large volume of resource have to be allocated to complete the project. One such project is a ship building in which it takes several years to complete it. The due date of such project is very important because the violation of the due date would result in a large amount of cost or penalty. In the planning of this kind of project, one of the most difficult considerations is various resource requirements over the entire period of the project. (for example, man power, work place, and facilities, etc.) It is not efficient to have all the resource for the project to resolve the fluctuation of the resource requirements. It can be better if some of the resources can be outsourced to optimize business operations. One of the other cases for such project is a software development in which a large volume of manpower is required. In this project, there are precedence considerations among the groups of jobs. Each group of jobs has its own due date to satisfy overall due date requirements. To meet the due date and promote cost competitiveness, outside resources with any economic price can be used. Therefore, the outsourcing strategy is one of the options that decision makers can consider to improve competitiveness in business.

Outsourcing has attracted a lot of interest in recent years. McCarthy and Anagnostou (2004) studied the rationale of outsourcing from an economic viewpoint. Cachon and Harker (2002) and Kaipia and Tanskanen (2003) investigated why many companies consider outsourcing as an attractive strategy to alleviate price competitiveness. Kim (2003) examined the problems of selecting outsourcing resources (companies). Kolisch (2000) addressed a make-to-order production problem with outsourcing, for which he has derived a mixed-integer programming model to minimize holding and setup costs of the associated supply chain. Lee et al. (2002) presented a model for an advanced planning and scheduling (APS) problem with outsourcing, for which they have derived a genetic algorithm to minimize the makespan. Lee and Sung (2008) suggested outsourcing scheduling problems with various objective functions. Various scheduling problems with outsourcing in different multi-stage production systems have been considered (Choi and Chung, 2011; Lee and Choi, 2011; Neto and Filho, 2011; Qi, 2011; Mokhtari et al., 2012). Mishra et al. (2008) developed a Tabu-simulated annealing method for a mixed-integer pro-
In the next section, the problem complexity and several optimal properties are developed.
Proposition 2. For the problem $\mathcal{I}\mid C_j \leq d_j \mid \sum_{k \in K} \sum_{j \in \mathcal{O}_k} o_{jk}$, there exists an optimal schedule in which all the non-outsourced (i.e., in-house) jobs are sequenced with no idle time.

Proof. In the optimal schedule $\pi^*$, the constraints $\sum_{i \in J, j \in \mathcal{O}_i} (p_i + o_j) + p_j \leq d_i$ will be satisfied. Then the schedule with $o_j = 0$ also satisfies the constraints and has the same objective function value.

This completes the proof.

Corollary 1. From Proposition 1 and 2, the problem can be reduced to both the selection of jobs to be outsourced and the assignment of the outsourced jobs to an outsourcing provider.

Proof. Corollary 1 can be easily proven, so the proof is omitted.

Let decision variable $y_{jk}$ represents if the job $j$ would be assigned to the outsourcing provider $k$. The variable $y_{jk}$ will have 1 if the job $j$ is assigned to the outsourcing provider $k$. Otherwise, $y_{jk}$ will be 0. Then, the integer programming model for the problem is as follows:

$$\min \sum_{k \in K} \sum_{j \in \mathcal{O}_k} o_{jk} y_{jk},$$

subject to

$$\sum_{i \in J} p_i (1 - \sum_{k \in K} y_{ik}) + p_j \leq d_j + M \sum_{k \in K} y_{jk}, \forall j \in J,$$

$$\sum_{j \in \mathcal{O}_k} y_{jk} \leq R_k, \forall k \in K,$$

$$\sum_{k \in K} y_{jk} \leq 1, \forall j \in J,$$

$$y_{jk} \in \{0, 1\}, \forall j \in J, k \in K,$$

where $M$ represents any big value.

III. SOLUTION PROCEDURE FOR THE PROBLEM

In this section, a network modeling of the problem is suggested and a search procedure in the network model is also considered to find the optimal solution with efficiency.

To make a network model of the problem, several definitions are needed. Let $\mathcal{O}(i)$ be a partial schedule for jobs 1, 2, ..., $i$, where $i \in J$. Using the definition of $\mathcal{O}(i)$, the problem $\mathcal{I}\mid C_j \leq d_j \mid \sum_{k \in K} \sum_{j \in \mathcal{O}_k} o_{jk}$ can be modeled as a network problem and solved by applying a search algorithm. For the network model of the problem, all the possible combinations of $(i, m, r_i, ..., r_k)$ denotes the nodes in the network where $m, r_i, ..., r_k$ are any integers satisfying $0 \leq m \leq \sum_{j \in \mathcal{O}_k} p_j$, $r_i \leq R_i$, ..., $r_k \leq R_k$. The network modeling of the problem starts with the root node, $(0, 0, 0, ..., 0)$. Links and nodes are added into the network at each iteration of the network modeling. From the root node, links and nodes are generated to make the whole network model of the problem. If we have a node $(i, m, r_i, ..., r_k)$ in the network model, then there can be 1 + $|k|$ links from this node such as $(i + 1, m + p_i, r_i, ..., r_k), (i + 1, m, r_1 + r_{i+1}(r_i), ..., r_k), (i + 1, m, r_1, ..., r_k + r_{i+1}(r_k))$ respectively. If the inequality $(m + p_i) \leq d_i$ is satisfied, then there is a link from $(i, m, r_i, ..., r_k)$ to $(i + 1, m + p_i, r_i, ..., r_k)$ and the weight of the link is the zero. If the inequality $r_k + r_{i+1}(r_k) \leq R_k$ is satisfied for $k \in K$, then there is a link from $(i, m, r_1, ..., r_k)$ to $(i + 1, m, r_1, ..., r_k + r_{i+1}(r_k))$ and the weight of the link is $o_{i+1}(r_k)$. The maximum number of links from each node to others is $|K| + 1$. Note that $r_k$ in the node $(i, m, r_1, ..., r_k)$ denotes the capacity that has been consumed at outsourcing provider $k$ at node $(i, m, r_1, ..., r_k)$.

From the following procedure, we can generate a network model for the problem $\mathcal{I}\mid C_j \leq d_j \mid \sum_{k \in K} \sum_{j \in \mathcal{O}_k} o_{jk}$.

Network generation procedure

1. Let a set of nodes $N_1$ and a set of links $L$ be empty, i.e. $N_1 = \emptyset$ and $L = \emptyset$. And let a set of nodes $N_2$ have a root node, initially, i.e. $N_2 = \{(0, 0, 0, ..., 0)\}$.
2. Select an arbitrary node in the set $N_2$ and let $(i, m, r_1, ..., r_k)$ be the selected node. If $N_2 = \{(0, 0, 0, ..., 0)\}$, then $(0, 0, 0, ..., 0)$ will be the selected node.
3. If the inequality $(m + p_i) \leq d_i$ is satisfied for the selected node, then go to step 4. Otherwise go to step 5.
4. The link from $(i, m, r_1, ..., r_k)$ to $(i + 1, m + p_i, r_i, ..., r_k)$ is added into the set of links $L$ with its weight of zero. The node $(i + 1, m + p_i, r_i, ..., r_k)$ is added to the set $N_2$.
5. Iterate steps 6 and 7 for all $k \in K$.
6. If the inequality $r_k + r_{i+1}(r_k) \leq R_k$ is satisfied for $k \in K$, then go to step 7. Otherwise, perform step 6 again with another $k \in K$.
7. The link from $(i, m, r_1, ..., r_k)$ to $(i + 1, m, r_1, ..., r_k + r_{i+1}(r_k), ..., r_k)$ is added into $L$ and its weight is $o_{i+1}(r_k)$. The node $(i + 1, m, r_1, ..., r_k + r_{i+1}(r_k), ..., r_k)$ is added into $N_2$.
8. If all the iterations of steps 6 and 7 for the selected node $(i, m, r_1, ..., r_k)$ are completed, then the node $(i, m, r_1, ..., r_k)$ is added into $N_2$ and subtracted from $N_2$.
9. If the set $N_2$ is not empty yet, go to step 2.

Fig. 1 is the flow diagram of the network generation procedure. Let $G(\mathcal{I}\mid C_j \leq d_j \mid \sum_{k \in K} \sum_{j \in \mathcal{O}_k} o_{jk})$ be the network generated with problem $\mathcal{I}\mid C_j \leq d_j \mid \sum_{k \in K} \sum_{j \in \mathcal{O}_k} o_{jk}$ by applying the procedure depicted in Fig. 1. Note that the network $G(\mathcal{I}\mid C_j \leq d_j \mid \sum_{k \in K} \sum_{j \in \mathcal{O}_k} o_{jk})$ takes the shape of a tree. In $G(\cdot)$, a path from the root node, $(0, 0, 0, ..., 0)$, to any end node constructs a feasible schedule and the sum of the link weights on the path is the sum of the outsourcing cost of the corresponding schedule. A solution algorithm that searches all the paths from the root node to the end nodes in the tree $G(\cdot)$ can be considered. The algorithm calculates all the shortest
paths to all of end nodes in the network and the minimum value among the shortest paths to all the end nodes is the optimal value of the problem. Due to the fact that the total number of paths in the tree \( G(\cdot) \) cannot exceed the value \( |K|! \), the optimal solution for problem \( \max |C_j| \leq d_j \) can be found in the exponential time complexity with the order of \( (|K|+1)! \) by applying a search algorithm. However, because of the exponential complexity of the search algorithm, some modifications are needed to improve the efficiency.

In the network modeling of the problem, all the end nodes of the network (tree) should be explored to find the optimal solution. To make this search procedure to be more effective, Depth First Search (DFS) procedure is suggested for the graph \( G(\{i|C_j \leq d_j \sum_{k<i} \sum_{j \neq k} o_{jk} \}) \). DFS starts with the root node, \((0, 0, 0, \ldots, 0)\). A node will be selected as a next search node among the nodes linked to \((0, 0, 0, \ldots, 0)\). At each iteration of DFS, the link from \((i, m, r_1, \ldots, r_k)\) to \((i+1, m+p_i, r_1, \ldots, r_k)\) has a higher priority than any other links when selecting a node for the next search. Let a solution \( S(DFS) \) be the first feasible schedule found by DFS and the set \( O(DFS) \) be the subset of jobs that are selected to be outsourced under the schedule \( S(DFS) \). The set \( NO(DFS) \) is the set of jobs that are selected to be processed in the in-house machine on the schedule \( S(DFS) \). We assume that all the jobs in the set \( J = O(DFS) \cup NO(DFS) \) are indexed in EDD order. \(|J|\) is the number of jobs in the set \( J \). Then the pseudo-code using the recursive function \( dfs(\cdot) \) for the DFS is as follows:

**Procedure DFS**

- set \( job.number = 0 \), \( best.solution = \text{MAX}, \)
- call function \( dfs(job.number) \)

**function dfs(job.number)**

- if \((job.number > |J|)\), then return
- if \( \text{('the total outsourcing cost so far') \ > \ best.solution} \), then return
- update \( \text{('the total outsourcing cost so far')} \) with \( \pi(job.number) \)
- update \( best.solution \),
- if \((job.mp + p_{job.number+1} \leq d_{job.number+1})\),
  - then update \( \pi(job.number + 1) \)
  - call function \( dfs(job.number + 1) \)
- for \((k = 1; k < |K|; \ k++)\)
  - if \((r^k_{\pi(job.number)} + r^k_{\pi(job.number+1)}) \leq R_k \),
    - then update \( \pi(job.number + 1) \)
    - call function \( dfs(job.number + 1) \)

In Procedure DFS, \( job.number \) denotes the current depth of the search and \( \text{MAX} \) represents any big value. The procedure DFS starts with \( dfs(1) \), where \( job.number \) equals 1. ‘the total outsourcing cost so far’ denotes the sum of the outsourcing cost from job 1 to \( job.number \). \( best.solution \) indicates the best solution found so far. \( r^k_{\pi(job.number)} \) denotes \( \sum_{j \neq k} r_{jk} \). Lastly, \( job.mp \) represents \( C_j \). Let \( \pi(\cdot) \) be the first feasible schedule found by applying Procedure DFS. Then some optimality conditions for the problem are suggested as in the following propositions.

**Proposition 3.** If \(|O_S| \leq 1 \) for the problem with a constant outsourcing cost \( o_k = o \) for \( \forall j \in J \) and \( \forall k \in K \) and with arbitrary due date \( d_j \) for \( \forall j \in J \), then the schedule \( \pi(\cdot) \) is the optimal solution.

**Proof:**

1. If \(|O_S| = 0 \), then the schedule \( \pi(\cdot) \) is the optimal.
2. In case of \(|O_S| = 1 \), let the job \( j \) be the element of the set \( O_S \), which means the condition \( \sum_{i \neq j} p_i > d_j \) is satisfied.

Then we have two cases at the optimal solution. The first one is at least one more jobs not in the outsourced group should be added to the outsourced group. The second is that job \( j \) should be interchanged with the precedent job \( i \), \( d_i < d_j \) to make a feasible schedule which has no outsourced job. In the first case in which at least one more
job should be outsourced, the schedule $\pi(\cdot)$ is optimal because $o_{jk} = 0$ for all $j \in J$ and $k \in K$. In the second case, if we assume that there is the feasible schedule $\pi^*$ of $|O| = 0$, then (from Proposition 1) there should exist an optimal schedule such that all the jobs are sequenced in EDD order, which contradicts that the schedule $\pi^*$ is feasible.

This completes the proof.

Proposition 4. For the problem with constant processing time $(p_j = p)$ for all $j \in J$ and with arbitrary due date $d_j$ for all $j \in J$, the schedule $\pi(\cdot)$ always gives the minimum number of outsourced jobs among all the feasible solutions.

Proof:

For the problem with constant processing time $(p_j = p)$ for all $j \in J$, the schedule $\pi(\cdot)$ in which all the jobs satisfying the condition $\sum_{i \in J, i \neq o_{jk}} p + p \leq d_j$ are in the set of non-outsourced jobs gives the schedule with the minimum number of outsourced jobs.

This completes the proof.

With arbitrary outsourcing cost, $o_{jk}$, the schedule with the minimum number of outsourced jobs does not always give the optimal schedule. In Proposition 5, the optimality condition for the schedule with the minimum number of outsourced jobs is presented.

Proposition 5. For the problem with constant processing time and outsourcing cost $(p_j = p, o_{jk} = o)$ for all $j \in J$ and $k \in K$ and with arbitrary due date $d_j$ for all $j \in J$, the schedule $\pi(\cdot)$ always gives the optimal solution.

Proof:

From Proposition 4, the schedule $\pi(\cdot)$ gives the minimum number of outsourced jobs among all the feasible solutions. Because $o_{jk} = 0$ for all $j \in J$ and $k \in K$, this gives the optimality conditions.

This completes the proof.

To compare the efficiency of Procedure DFS, a DP (Dynamic Programming) algorithm, which is known to be more efficient than a general search algorithm, is derived for the problem. Let a value function $f(i, m, r_1, ..., r_K)$ be the minimum length of the path from the root node to a node $(i, m, r_1, ..., r_K)$ and a set $V(G(\cdot))$ be the set of nodes in network $G(\cdot)$. Then the following recurrence relation can be derived. The value function $f(i, m, r_1, ..., r_K)$ is a minimum value among $f(i-1, m-p_i, r_1, ..., r_K)$ and $f(i-1, m, r_1, ..., r_{-1}, r_{+1}, ..., r_K) + o_{jk}$ for all $k \in K$ and $(i-1, m, r_1, ..., r_{-1}, r_{+1}, ..., r_K) \in V(G(\cdot))$.

The optimal value is $\min_{i, m \in J} \{ f(i, m, r_1, ..., r_K) \}$ where $(i, m, r_1, ..., r_K) \in V(G(\cdot))$. The DP algorithm for the problem $\exists j : d_j \leq d$ is summarized as follows:

Dynamic Programming DP
1. Set $f(j, m, R_1, ..., R_K) = 0$ for all non-negative integers $m$ less than or equal to $min \{ d_{ij} \} \sum_{j \in J} p_j$.
2. For the $i$-th iteration, the value function $f(i, m, r_1, ..., r_K)$ can be calculated as follows:

$$f(i, m, r_1, ..., r_K) = \begin{cases} f(i-1, m-p_i, r_1, ..., r_K), & \text{if } m-p_i \geq 0 \text{ and } m \leq d_j, \vspace{1mm} \\ f(i-1, m, r_1, ..., r_K) + o_{jk}, & \text{if } r_j - r_{jk} \geq 0, \vspace{1mm} \\ \infty, & \text{otherwise}. \end{cases}$$

Because the total number of states of the Procedure DFS cannot exceed the value $|I| \times \sum_{j \in J} p_j \times R_1 \times ... \times R_K$, an optimal solution for the problem can be found in the pseudo-polynomial complexity order of $O(|I| \times \sum_{j \in J} p_j \times R_1 \times ... \times R_K)$. The complexity of the DP algorithm is the same as that of Procedure DFS. Two solution procedures, Procedure DFS and DP algorithm, have the same complexity because these two procedures search the optimal solution based on the same network model. However, the efficiency of them would differ because they search the optimal solution with different search methods. The efficiency of these two solution procedures are tested and compared to each other in the next section.

IV. NUMERICAL EXPERIMENTS

In this section, the performances of the suggested procedure are evaluated. Computational results of the algorithm are presented with randomly generated problems. The parameters of the problems are generated with the predefined statistical distributions. The statistical distributions of the test instances are selected in the similar way as in the references (Koulamas, 1994; Choi et al., 2005; Lee and Sung, 2008). The processing time $p_j$ and the outsourcing cost $o_{jk}$ are generated from the uniform distribution between 1 and 10, and between 1 and 30, respectively. The due date $d_j$ is randomly generated from a uniform distribution depending on the sum of the processing time and two parameters Span of Due Date (SDD) and Tardiness Factor (TF). SDD denotes the degree of span of due dates, which means that the larger value of SDD will result in the wider range of $d_j$'s value. TF denotes the tardiness factor which is approximately proportional to the ratio of jobs likely to be tardy (Raman and Talbot, 1993). We assume that $d_j$ has a uniform distribution between $\frac{p_{\text{sum}}}{2} \times (1-TF+\text{SDD/2})$ and $\frac{p_{\text{sum}}}{2} \times (1-\text{TF+\text{SDD/2}})$, where $p_{\text{sum}} = \sum_{j \in J} p_j$. The difficulties of the problems may be determined by the values of SDD and TF (Lee and Sung, 2008). In our experiments, SDD and TF are selected among 0.2, 0.4, 0.6, 0.8, and 1.0. For each combination of SDD and TF, 30 instances are generated. The capacity
of \( k \)-th external facility, i.e., \( R_k \), is set to \( CF \times \sum_{j \in J} f_{jk} \), where \( CF \) is the capacity factor that is selected among 0.2, 0.4, 0.6, 0.8, and 1.0.

For the numerical experiments, C language was used for implementing the suggested algorithms. The machine with Core i-5 CPU with 4G memory was used for running the numerical experiments. In Table 1, the experimental results for Procedure DFS (or DFS algorithm) and Dynamic Programming DP (or DP) are presented. Computational times are measured in seconds.

From Table 1, we can show that the DFS algorithm has better results than DP in efficiency; the DFS algorithm can solve larger sized problems. In Table 1, it is noted that the efficiency of the suggested algorithms are significantly affected by the number of jobs and the degree of SDD, but relatively less affected by the TF.

The computational results in Table 2 show that the number of jobs in the problem considerably affects the efficiency of DP, but not really affects the suggested algorithm, DFS. From Tables 1 and 2, it is noted that the outsourcing cost (i.e. the objective function value of the problem) is more affected by SDD than TF.

For each \(|J|\) and \(|K|\) values, various SDD and TF values are tested as in Tables 1 and 2. Table 3 represents the computational results of DFS algorithm for larger numbers of jobs and outsourcing providers. The computational times in column b and the minimum objective function values in column c in Table 3 are the average values of the results executed with various SDD and TF values.

In Table 3, it is observed that almost all of the computational times of DFS algorithm are reasonably short even for large numbers of jobs and outsourcing providers. The efficiency of the suggested algorithm are affected by the both the number of jobs and the number of outsourcing providers.

As shown in Table 3, it is reasonable to assume that the cost will be decreased as the number of alternatives in the problem increases. For example, the minimum objective function value with \(|J| = 30\) and \(|K| = 2\) is 34.4448 in contrast to 27.1571 in case of \(|J| = 30\) and \(|K| = 3\). It is noted that the outsourcing cost (i.e. the objective function value of the problem) decreases as the number of outsourcing providers is increased. The same can be applied to the real world problem. If we have more choice for the outsourcing providers, the cost will be decreased. That is the motivation for us to consider the scheduling problem with multiple outsourcing providers.
developed the DSF algorithm efficiently searching the network properties have been presented. To solve the problem, we have developed the DSF algorithm efficiently searching the network converted from the problem.

The computational experiments of the proposed scheme have been performed for randomly generated problems. The proposed procedure has illustrated outstanding performance in comparison to DP, even for larger numbers of jobs and outsourcing providers.

The outsourcing is one of the business strategies to improve the competitiveness of the company. As the competition increases in the global economy, the importance of the outsourcing strategy will grow. The scheduling problem with outsourcing and also with additional real world constraints can be considered for further research subjects. For example, the logistics cost constraints for the outsourcing or the precedence constraints among jobs can be addressed as a next step. Non-zero arrival times of jobs can also be considered in the future work.

**REFERENCES**


**Table 3. Computational results for DFS algorithm with various |J| and |K| values.**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>Computational time of DFS</th>
<th>Minimum objective function value</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>2</td>
<td>0.0000</td>
<td>14.1595</td>
</tr>
<tr>
<td>15</td>
<td>2</td>
<td>0.0001</td>
<td>19.3097</td>
</tr>
<tr>
<td>20</td>
<td>2</td>
<td>0.0041</td>
<td>26.6097</td>
</tr>
<tr>
<td>30</td>
<td>2</td>
<td>0.2582</td>
<td>34.4448</td>
</tr>
<tr>
<td>20</td>
<td>3</td>
<td>0.0033</td>
<td>19.2985</td>
</tr>
<tr>
<td>30</td>
<td>3</td>
<td>1.6955</td>
<td>27.1571</td>
</tr>
<tr>
<td>40</td>
<td>3</td>
<td>108.8124</td>
<td>35.5917</td>
</tr>
</tbody>
</table>

**V. CONCLUSIONS**

In this paper, we have studied the single machine scheduling problem with multiple external facilities. The problem is to minimize the total outsourcing cost subject to the outsourcing capacity and the due date constraints. Several optimality properties have been presented. To solve the problem, we have developed the DSF algorithm efficiently searching the network converted from the problem.

The computational experiments of the proposed scheme have been performed for randomly generated problems. The proposed procedure has illustrated outstanding performance in comparison to DP, even for larger numbers of jobs and outsourcing providers.

The outsourcing is one of the business strategies to improve the competitiveness of the company. As the competition increases in the global economy, the importance of the outsourcing strategy will grow. The scheduling problem with outsourcing and also with additional real world constraints can be considered for further research subjects. For example, the logistics cost constraints for the outsourcing or the precedence constraints among jobs can be addressed as a next step. Non-zero arrival times of jobs can also be considered in the future work.