Mathematical Models for Optimization Problem in Flowshop Scheduling

Jun Wang*
School of Economics and Management, Beihang University, Beijing 100191, China
*Corresponding author (E-mail: wangjun23@139.com)

Lingdi Liu
School of Economics and Management, Beihang University, Beijing 100191, China

Jing Wang
School of Economics and Management, Beihang University, Beijing 100191, China

Abstract
This paper particularly refers to a tobacco primary flowshop and describes a modeling approach for flowshop scheduling problem. In the tobacco primary flowshop, the facilities which contribute most to the utilization of raw materials are usually the cut tobacco line. Since the cost differs for transferring between different tobacco batches, the cost is the bottleneck in terms of scheduling. Thus, in order to save on costs, a planning approach which aims at minimizing the number of transferring throughout the horizon is needed besides optimizing the process design in the tobacco primary flowshop. However, Flowshop Sequencing Problem (FSP) is NP-complete problem and hard to model directly. Rather than adopting a mathematical model for developing a comprehensive production plan, a Traveling Salesman Problem (TSP) model is preferred. The tobacco batches are regarded as cities and the optimal permutation of producing cut tobacco might be the optimal tour. As well, the flowshop scheduling problem without time restraint can be modeled as Classical Traveling Salesman Problem (CTSP) and the one with time restraint can be modeled as Traveling Salesman Problem with Time Windows (TSPTW). Then, we use Matlab to provide a producing permutation of the batches. We can indicate that modeling flowshop scheduling problem into TSP is efficient. Finally, we suggest some interesting directions for future researches.

Key words: Flowshop sequencing problem, Traveling salesman problem, Tobacco primary flowshop, Time restraint, Genetic algorithm

1. INTRODUCTION
In material intensive industries such as cigarette industry, clothes industry, and sugar industry, the bottleneck is to arrange the producing permutation in the primary flowshop. The raw material productions have to satisfy the followed process requirement. Here, we particularly consider the tobacco primary flowshop. In the flowshop, all jobs follow the same machine sequence and each job has exactly one operation on each machine (Bagchi, Gupta and Sriskandarajah, 2006). Flowshop scheduling problem is to find an optimal sequence of producing the jobs so that the cost is minimized. Flowshop scheduling problem is always a family scheduling model, where jobs are partitioned into families according to their similarity, so that no setup time is required for a job which belongs to the same family of the previously processed job. However, a leadtime is required at the beginning of the schedule and on each occasion when the machine switches from processing jobs in one family to jobs in another family (Potts and Kovalyov, 2000). As we assumed, a job family stands as a tobacco brand and includes several batches. The cost of switching from processing batches of a brand to those in other brands is higher than the cost of switching in the same brand. We need build different models to fit different situations.

As we can see, flowshop scheduling problem is NP-completed hard (Garey, Johnson and Sethi, 1976, Gonzalez T. and Sahni, 1978), which means that solving flowshop scheduling problem directly will cost exponential time when the data quantity increases. Thus, a method which can effectively model the process and find an optimal permutation of batches need be stated. In some articles, flowshop scheduling problem can be modeled as traveling salesman problems (TSP) and then solved by using efficient algorithms available for the TSP (Wang, 2012, Al-Dulaimi and Ali, 2008). The classic traveling salesman problem (CTSP) is to find the shortest tour which the salesman visits each city exactly once on a giving map. As time flies, TSP is improved into different models such as asymmetrical TSP (ATSP), which has different length when salesman go to city j from city i and go to i from j; TSP with pickup and delivery (TSPPD); TSP with time windows (TSTPW), which means the salesman should arrive each city in time; etc. Since TSP is a binary integer optimization problem, bound-and-batch method can be used to solve TSP (Bagchi, Gupta and Sriskandarajah, 2006). However, TSP is a NP-complete problem, which means that the time consumption increased exponentially when the...
number of cities increases. There are also many heuristic algorithms can be used to solve the TSP such as Hopfield neural network, simulated annealing and genetic algorithm (Wang, 2012, Al-Dulaimi and Ali, 2008). Thus we can model flowshop scheduling problem into TSP and use the developed methods of solving TSP to solve flowshop scheduling problem.

Our research aims to develop an effective method capable of modeling and solving flowshop scheduling problem. Our procedure is reformulating flowshop scheduling problem into TSP and solving the problem by using some efficient algorithms.

This paper is organized as follows: Part II illustrates the tobacco primary flowshop in a two-stage cigarette production to give a general realization of flowshop scheduling problem. Then, in part III, to solve scheduling problem in the tobacco primary flowshop, a graph whose vertex set consists the batches and edge set consists the transferring costs is formulated so that flowshop scheduling problem can be modeled into TSP. Next, by using some usual transformations, the different models are formulated respectively. The flowshop scheduling problem under time restraint can be rewritten as TSPTW (Allahverdi, Ng, Cheng and Kovalyov, 2008). At last, the software Matlab is used to assist our calculation and give a data analysis in part IV.

2. ENVIRONMENT ANALYSIS

2.1. Cigarette Manufacturing

Cigarette production process has many characteristics such as multi stages, multi varieties and mandatory plan, and it is the typical hybrid production mode, there is an urgent demand for production planning optimization. Figure 1 shows the procedures of a two-stage production process of cigarette. The production is divided into two main stages, the primary stage is the cut tobacco manufacturing, and the secondary stage is cigarette packing. The cut tobaccos produced by the primary stage are transferred to the secondary stage with parallel feeders. Therefore, the production quantity and delivery in the primary stage is controlled by the requirements of the secondary stage. And the permutation of producing batches must be arranged to satisfy the leadtime requirement in the secondary stage.

![Figure 1. Two-stage Production Process of Cigarette](image)

Generally, a tobacco primary flowshop gain cut tobacco by several processes including: a step of blade pretreatment; a step of cut tobacco processing; and a step of storing the cut tobacco which the alcohol has been added and the content of flavor-enhancing esters is increased (Ito and Kenji, 1985, Xia and Peng, 2009). The processes are exactly consisted in the main assembly line in fig. 1. In the fig.1, there is a main assembly line cooperated with two subordinate lines, stem line and inflate line.

![Figure 2. Sketch map for tobacco primary flowshop](image)
2.2. Products Characteristics

For each brand, cut tobacco are gained by producing raw materials under a fixed proportion, so the materials cost and process time are the same. But that of different brand is different (Ito and Kenji, 1985). According to the capacity and technique limitation, the requirement of a brand is always divided into several batches in practical manufacturing. As the need of homogenization, each batches is generally fixed formula feeding, and organized according to the requirements of strict rules of process. But the parallel unit production in packing phase makes it difficult to producing continuously of cut tobacco in the same brand. The transferring cost in same brand of cut tobacco is lower than the cost of switching between different types of cut tobacco batch. In the process of production, switching cost between different types of tobacco is divided into two cases: switching from high-qualified brands to low-qualified brand, which need not clean production line and the cost is low; on the other hand, the switching from low-end brands to the high-end brand, which need to clean the production line, so the cost is high.

2.3. Optimization Objective

Tobacco primary flowshop is the first stage in the cigarette manufacturing system and the finished materials are transferred to second flowshop to satisfy the producing requirement. The optimal objective in this stage is to allocate a permutation of different batches so that the switching cost is minimized (Liu, Chen, Cui and He, 2003). General, all the finished materials in the tobacco primary flowshop are produced in the same process route and present the typical process plant characteristics. Once the finished material feeding order is determined, the production in subordinate line will be affected according to the formula of blending, whose schedule follows the arrangement of the main assembly line. Organize production is according to the semi-finished products inventory.

3. MODEL ANALYSIS

The jobs in a tobacco primary flowshop are divided into \( n \) batches which will be processed on a single main assembly line. There is a cost for transferring between batches. And each batch has a process time on the assembly line. Our target is to minimize the total cost and make sure that each batch can be finished before the required time.

According to our consideration, we define a weighted graph \( G = (Prodw, ct, ProdT, ProdD) \) with vertex set \( Prod = \{1,2,\ldots,n\} \), whose elements are the batches, which can be seen as cities to be visited; edge set \( ct = \{ct_{ij}|i,j \in Prod\} \), whose elements are the cost of transferring from batch \( i \) to batch \( j \), which is likely the traveling distance between city \( i \) and city \( j \); time restraint set \( ProdT = \{L_i|i \in Prod\} \), in which \( L_i \) illustrates the latest time required to finish manufacturing batch \( i \); and duration times set \( ProdD = \{dt_i|i \in Prod\} \), in which \( dt_i \) means the process time for manufacturing batch \( i \), which can be seen as the duration time at city \( i \). In our assumption, \( G \) is an asymmetric graph, which means that \( ct_{ij} = ct_{ji}, i,j \in Prod \).

Then we can derive a model which formulates flowshop scheduling problem into TSP:

Objective: \[
\begin{aligned}
&\min \sum_{i=1}^{n} ct_{ij} z_{ij} + M \cdot \max \left( \left[ \text{lot}T_i + \sum_{j \in S_{\mathcal{P}}} dt_{nj} \right] - L_i, 0 \right) \\
\text{s.t.}
&\sum_{j=1}^{n} z_{ij} = 1, \forall i \in Prod \\
&\sum_{i=1}^{n} z_{ij} = 1, \forall j \in Prod \\
&\sum_{i \in \mathcal{SPR}} \sum_{j \in \mathcal{SPR}} z_{ij} \leq |\mathcal{SPR}| - 1, \forall \mathcal{SPR} \subseteq Prod \\
&2 \leq |\mathcal{SPR}| \leq n - 1
\end{aligned}
\]
\[ z_{ij} \in \{0,1\} \] (6)

The parameters and variables are defined as follows:
- \( z_{ij} \): binary variable indicating that travelling from vertex \( i \) to vertex \( j \)
- \( c_{tij} \): the cost of transferring from vertex \( i \) to vertex \( j \)
- \( M \): a large number
- \( n_j \): the order of vertex \( j \) in the final permutation of vertices in \( Prod \)
- \( lotT_i \): the actual time of arriving vertex \( i \)
- \( dt_{nj} \): the duration time at vertex \( n_j \)
- \( L_i \): the latest requirement time to visit vertex \( i \)
- \( SPR \): subsets of vertex set \( Prod \)

Where \( z_{ij} = 1, i, j \in Prod \) means travelling from vertex \( i \) to vertex \( j \), and \( z_{ij} = 0, i, j \in Prod \) means the doesn’t travelling from vertex \( i \) to vertex \( j \), which also shows the transferring states between batch \( i \) and \( j \) in active production; \( SPR \) can be any subsets of vertex set \( Prod \); and when \( n_j \leq n_i, n_j \in Prod \), vertex \( j \) is visited before visiting vertex \( i \) which means brand batch \( n_j \) is manufactured before batch \( n_i \) in active production.

Equation (1) is the objective and wants to minimize the cost; where the first part \( \sum_{i=1}^{n} c_{tij} z_{ij} \) means the makespan of visiting all of the vertices in \( Prod \); the second part \( M \cdot \sum_{i=1}^{n} \max\left(\left(\text{lot}T_i + \sum_{j \in SPR} dt_{nj}\right) - L_i, 0\right) \) is the time restraint, which is 0 if the machine finishes manufacturing batch \( i \) before the latest time required to finish manufacturing batch \( i \), i.e. \( \text{lot}T_i + \sum_{j \in SPR} dt_{nj} < L_i \); is positive when the machine doesn’t finish manufacturing batch \( i \) before the latest time required to finish manufacturing batch \( i \), i.e. \( \text{lot}T_i + \sum_{j \in SPR} dt_{nj} < L_i \); and in the second case, the product \( M \cdot \sum_{i=1}^{n} \max\left(\left(\text{lot}T_i + \sum_{j \in SPR} dt_{nj}\right) - L_i, 0\right) \) is a very large positive number since we set \( M \) as a very large number. Therefore, the second part of objective function must be zero in the optimal solution so that each batch can be manufactured before the latest time required finishing manufacturing. Equations (2) and (3) are assignment constraint, which mean that we will go through each vertices and the numbers of input and output acre can only be 1. Equations (4) and (5) ensure that the feasible solution doesn’t include sub cycle, i.e. the batch sequence is continuously arranged. Equation (6) shows the value set of \( z_{ij} \).

To solve this model, we need find a Hamilton cycle for which the cost of the edge set is minimized. A Hamilton cycle of graph \( G \) is a cycle passing through each vertex exactly once (Cornuejols, Naddef, and Pulleyblank, 1985).

Since the model (1)-(6) is non-linear, we dissolve the model under two different situations: the model without time limitation and the model under time limitation. We analyze the two situations in the following two subsections. For each situation, the model is reformulated on basis of model (1)-(6).

### 3.1. The model without time restraint

The graph \( G_1 = (Prod, ct) \) has the same vertex set and edge set as graph \( G \) and lack time limitation set and duration restraint set by default. We introduce a variable \( u_i \in \{1,2,\cdots,\text{LOTN} - 1\}, i = 2,3,\cdots,n, \) to represent the number of finished batches before manufacturing batch \( i \). Then the model (1)-(8) can be reformulated as follows:

**Obj.**

\[ \min \sum_{i=1}^{n} c_{tij} z_{ij} \] (7)

s.t.

\[ \sum_{j=1}^{n} z_{ij} = 1, \quad \forall i \in Prod \] (8)

\[ \sum_{i=1}^{n} z_{ij} = 1, \quad \forall j \in Prod \] (9)

\[ u_j \geq u_i + 1 - n \cdot \left(1 - z_{ij}\right) \quad j \geq i \] (10)

\[ u_i \in \{1,2,\cdots,\text{LOTN} - 1\}, \quad i = 2,3,\cdots,n \] (11)
\[ z_{ij} \in \{0,1\}, \quad i,j \in \text{Prod} \]  
\[ u_1 = 0 \]  

When the switching cost is asymmetric, the cost matrix \( c_t \) is asymmetric. Therefore, using the same model above, we can obtain an asymmetric TSP.

### 3.2. The Model under Time Limitation

For graph \( G_2 = (\text{Prod}, c_t, \text{ProdT}) \), we introduce variable \( y_i = \max\{\text{lotT}_i - L_i, 0\} \), which means to manufacture the \( i \)-th batch in time, and reformulate model (1)-(6) as the following:

**Obj**

\[
\min \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij}Z_{ij} + M \sum_{i=1}^{n} y_i
\]  

s.t.

\[
\sum_{j=1}^{n} z_{ij} = 1, \quad \forall i \in \text{Prod} \tag{15}
\]

\[
\sum_{i=1}^{n} z_{ij} = 1, \quad \forall j \in \text{Prod} \tag{16}
\]

\[
u_j \geq u_i + 1 - n \cdot (1 - z_{ij}), \quad j \geq i \tag{17}
\]

\[
u_i \in \{1,2,\cdots,\text{LOTN} - 1\}, \quad i = 2,3,\cdots,n \tag{18}
\]

\[
z_{ij} \in \{0,1\}, \quad i,j \in \text{Prod} \tag{19}
\]

\[
y_i \geq \text{lotT}_i - L_i \tag{20}
\]

\[
y_i \geq 0 \tag{21}
\]

\[
u_1 = 0 \tag{22}
\]

### 4. Algorithm Analysis

In front sections, we have introduced the approach of modeling flowshop scheduling problem into TSP and built several mathematical models based on actual production. The TSP is well known to be NP-hard for general models. We may use Genetic Algorithm to solve the models and use software Matlab to analyze our models and give an acceptable numerical solution, which is the permutation of job batches.

During the processes of genetic algorithm, some key points must be paid attention on:

i. **Population initialization.** We code the batches in random real number sequence. If the number of batches is \( n \), then the sequence length is \( n - 1 \) and each batch appears once. Considering the actual production environment, we set the initial index as population size \( M = 200 \), crossover probability \( P_c = 0.9 \) and mutation probability \( P_m = 0.05 \).

ii. **Fitness function.** The optimal of TSP is minimizing the cost, thus we can use optimal function as the fitness function directly.

iii. **Termination criterion.** We set following termination criterions: (1) The maximize iteration generation is 500; (2) If the ratio of average fitness of chromosomes in a population with contemporary best chromosome fitness greater than 0.9, the algorithm terminates; (3) If the best chromosome can maintain over 10 generations, the algorithm terminates.

### 5. Data Results

Some actual data are given in this section. There is a tobacco primary flowshop which has a main assembly line and without considering subordinate lines by default. The capacity of the main assembly line is 6000kg/hour. The produced cut tobaccos are in four brands named A, B, C, and D which divided into 9 batches respectively. Each brand can be produced by 6000kg at each batch.
Table 1 indicates the transferring time between brand \( i \) and brand \( j \) where the unit is minute. The elements are the \( c_{ij} \) in the model. We can easily see that the time cost in transferring batches within the same brand is less than that between different brands.

**Table 1. Time cost between brand \( i \) and brand \( j \)**

<table>
<thead>
<tr>
<th>brand</th>
<th>( A )</th>
<th>( B )</th>
<th>( C )</th>
<th>( D )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( A )</td>
<td>5</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>( B )</td>
<td>30</td>
<td>5</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>( C )</td>
<td>30</td>
<td>30</td>
<td>5</td>
<td>30</td>
</tr>
<tr>
<td>( D )</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>5</td>
</tr>
</tbody>
</table>

**Table 2. Required lead time for each brand**

<table>
<thead>
<tr>
<th>brand</th>
<th>batches</th>
</tr>
</thead>
<tbody>
<tr>
<td>( A )</td>
<td>1 480 960 /</td>
</tr>
<tr>
<td>( B )</td>
<td>0 360 720 1080</td>
</tr>
<tr>
<td>( C )</td>
<td>0 / / /</td>
</tr>
<tr>
<td>( D )</td>
<td>0 / / /</td>
</tr>
</tbody>
</table>

For actual production, we need set lead times so that the finished tobacco can be used in the next stage, the tobacco packing stage, in time. Table 2 indicates the required leadtime for the batches of each brand. The row of A, B, C and D is the brand row. The column of 1, 2, 3 and 4 is the batches column. There are three batches of brand A, four batches of brand B, a batch of brand C and a batch of brand D. The cross position \((j,i), j = A,B,C and D, i = 1,2,3 and 4\) is the latest requirement leadtime, whose unit is minute, for finishing producing the batch \( i \) of brand \( j \). For example, the element in the position \((A,1)\) means that the first batch of brand A requires to be finished producing at 0, which is the time for starting the secondary stage. Another example is that the position \((B,2)\) means that the second batch of the brand B requires being finished at 360 minutes.

Both of the model (7)-(12) and the model (14)-(22) use the data in Tab.1. The data in Tab.2 are only be used in model (14)-(22) as \( L_{q_j} \). After computed on Matlab system which is operated in the environment where the personal computer with 8.00GB RAM is installed the 64types windows7 operation system. We get the outcomes in less than 1 second. When the problem scale becomes larger, the time consumption need be further studied. All of the outcomes are shown in Table 3 and Table 4. And we dare say that modeling flowshop scheduling problem as TSP is an effective way in which a feasible permutation of batches is given in the tobacco primary flowshop.

**Table 3. The batch sequence without time restraint**

<table>
<thead>
<tr>
<th>permutation</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>serial number</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>8</td>
<td>7</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>brand</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>C</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>D</td>
<td></td>
</tr>
</tbody>
</table>

Tab.3 shows the permutation without time restraint. In the permutation, the batches 1 to 9 are ordered as 1, 3, 2, 8, 7, 4, 5, 6, and 9. That is to say, the assembly line produces batch 1 in brand A firstly, then batch 3 in brand A, and then batch 2 in brand A, etc. The cost of this permutation is 115 minutes. We notice that the batches in the same brand are allocated in successive order so that the transferring cost is less than switching between different brands.

When considering the time restraint and producing time of each brand, the outcome is shown in Tab.4, where we set the start time of producing tobacco as \(-640\) minutes. The batches are ordered as 1, 3, 2, 8, 9, 7, 4, 5, and 6 and the actual finished time (minutes) are \(-420, -355, -290, -200, -110, -20, 45, 110\) and 175, respectively. The transferring time cost is 115 minutes. We can find that the batches in the same brand are produced as many as possible until the time limitation works. One obvious example of this situation is that the serial 1, 2, and 3 in brand A are allocated successively in the order 1, 2, and 3. But when comparing with the permutation in Tab.3 and that in Tab.4, we can easily find that brand D is in the last position and the fifth position respectively. That is to say, when time restraint works, serial 9 in brand D is produced before producing the batches of brand B. This is thoughtful because the lead time requirements force us to produce as many products in different brands as possible before the starting time of the secondary serial rather than to produce with least transferring time.
Table 4. The batch sequence under time restraint

<table>
<thead>
<tr>
<th>permutation</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>serial number</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>C</td>
<td>D</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>required leadtime (mins)</td>
<td>0</td>
<td>480</td>
<td>960</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>360</td>
<td>720</td>
<td>1080</td>
</tr>
<tr>
<td>actual leadtime (mins)</td>
<td>-420</td>
<td>-355</td>
<td>-290</td>
<td>-200</td>
<td>-110</td>
<td>-20</td>
<td>45</td>
<td>110</td>
<td>175</td>
</tr>
</tbody>
</table>

6. CONCLUSION

Flowshop scheduling problem is one of the key problems in the production and operation of a wide range of applications. This paper gives an effective method of modeling and solving the flowshop scheduling Problem in tobacco primary flowshop. After a review of several researches and a brief introduction of the cigarette producing environment, the flowshop scheduling problem is modeled into the Travelling Salesman Problem (TSP). Then the TSP without time restraint is shown as model (7)-(13) and another under time restraint is shown as model (14)-(22). We formulate the genetic algorithm frame to solve our model. At the last part, an actual example is given to test our model. The software Matlab is used in solving the models. The Matlab outcomes are analyzed which show that modeling flowshop scheduling problem into TSP is effective to obtain the reasonable batch permutation.

In the subsequent studies have mainly been focused on two aspects, one is to multiple scene model, the rapid convergence algorithm study; second is the model analysis of the special requirements to deal with emergency replenishment. In this case the problem size increases, presenting the multi-dimensional characteristics.

Acknowledgements

Thank all participants in Beihang University. Their support is gratefully acknowledged. This research is partially supported by Chinese National Foundation of Natural Science (No.71172016; No. 71332003).

REFERENCES


