Fuzzy Cell Genetic Algorithm Approach for Flexible Flow-Line Scheduling Model

Arash Nasrolahi Shirazi*, Meghan Steinhaus†, Matthew Agostinelli*, Manbir Sodhi*

* Department of Mechanical, Industrial, and Systems Engineering
University of Rhode Island
Kingston, RI 02881
arashshirazi@uri.edu

†Department of Mathematics, US Coast Guard Academy,
New London, CT 06320

Abstract—This paper focuses on makespan minimization for the flow line scheduling problem using a Fuzzy Cell Genetic Algorithm (FCGA). Real world applications of this problem are commonly found in printing and electronic circuit board manufacturing industries. A generalized integer programming (IP) model for this problem is proposed. The Fuzzy Cell Genetic Algorithm (FCGA) is proposed to solve the IP model, which has been proven to be NP-hard. Sample problems are generated with known good solutions to evaluate the effectiveness of the FCGA approach. The FCGA matches the performance of the IP model for small sized problem instances and it is proven to be effective for larger problem instances.

I. INTRODUCTION

The flow line sequencing problem is defined by a set of jobs, and each job follows the same route from one machine to another throughout a manufacturing system. The sequence problem is determine the position of each job relative to other jobs for processing. The general form considered here has been derived from traditional manufacturing systems where each stage is comprised of one machine that all jobs must pass through. The expanded version of flow line sequencing is this flexible flow line sequencing problem, where each stage consists of more than one machine in parallel for all stages of operation to allow for concurrent processing in each stage. The flexible flow line is synonymous with a hybrid flow shop except that only one stage needs to feature parallel machines in a hybrid flow shop where the flexible flow shop considers parallel machines in all stages. Additionally, hybrid flow shops allow jobs to skip stages where flexible flow shops require jobs to visit all stages in order[1, 2].

Flexible flow lines have been used to model many real world environments including the production of circuit boards, automobiles, paper, textiles, and concrete[3–8]. Formally, flexible flow line systems can be defined by a set of \( n \) jobs to be processed in \( m \) stages with \( p \) number of parallel machines in each stage. Generally, flexible flow line problems have the following features in common:

1) The number of stages to perform the jobs must be greater than one.
2) The total number of stages must be less than the total number of machines.
3) All jobs must be processed in every stage but it is optional which machine processes the job in that stage.
4) Each job has a processing time represented by \( P_{t,s} \), where \( t \) is the index of job to process and \( s \) is the \( s^{th} \) stage.

In the general form of the flexible flow line problem, all machines in each stage are ready to perform the jobs at time zero, the parallel machines in each stage are identical, each machine can perform just one job at a time, and any job can be processed by any machine in a given stage. Setup time is negligible in the general form and buffers are infinite between stages. Figure 1 provides a simplified graphical form of the flexible flow line. The jobs are indexed by \( t \), the stages notated by \( s \) and each machine notated by \( k \) in each work center to process the jobs.

![Fig. 1. Graphical representation of the problem design which \( j \), \( s \) and \( p \) are index representative of jobs, stages and machines respectively.](image)

Generally, the objective of flexible flow line optimization is to minimize the makespan as a scheduling criteria. The makespan is defined as the time between the start of the first job in the system and the completion of the last job. Minimizing the makespan reduces the maximum completion time where all jobs are considered. This means each job has a completion time in the last stage and the maximum of all completion times is the makespan for the system. Traditionally, makespan minimization has been the focus of flexible flow
Several approaches have been taken to solve the traditional flow shop, parallel process shops, and hybrid flow shop problems. These include exact methods, heuristics, and meta-heuristics. However, there has been limited studies focusing specifically on the flexible flow line.

For exact methods, branch and bound algorithms have proven to be the preferred technique for solving flow line sequencing problems to optimality. The majority of branch and bound techniques have focused on solving scheduling problems with parallel machines [10–20] and hybrid flow shops [21–25]. In 1993, Sawik [26] first proposed a mathematical model for solving small-sized scheduling problems in a flexible flow line with limited buffers. Jungwattanakita et al. [27] proposed a mixed integer programming model for minimizing the makespan and the number of tardy jobs for a flexible flow line. In Jungwattanakita et al. [27] model, they considered job sequence and machine-dependent set up times.

The NP-hardness of the flexible flow line problem underscores the need and motivation to consider heuristic and meta-heuristic approaches [28]. Brah and Loo [29] evaluated the effectiveness of various heuristic methods that had not been applied to flow shops with multiple processors. There is limited literature that applies heuristic and meta-heuristic algorithms to flexible flow line problems. The proposed random keys genetic algorithm of Kurz and Askin [30] was applied to a flexible flow line model. They minimized makespan for a system with sequence dependent setup times. Akrami et al. [31] proposed GA and Tabu search heuristics for solving a specific flexible flow line sequencing model with limited intermediate buffers and compared results with an optimal enumeration method (OEM) on randomly generated problems. In Akrami et al. [31] flexible flow line model, sequencing, lot size, and product specific scheduling decisions are considered. They minimize total costs made up of the sum of setup costs, work-in-process inventory holding costs, and finished products inventory holding costs. Computational results from generated problems illustrated the promising performance of GA and Tabu search in comparison with OEM on large-size problems. Jenabi et al. [32] proposed a Hybrid Genetic Algorithm (HGA) and Simulated Annealing (SA) to solve flexible flow line problems with unrelated parallel processors on finite planning horizons. The main objective is to minimize the total setup and inventory holding costs per unit time without any stock-out. Final results proved the outperformance of HGA on generated problems compared with SA without considering the computation time. Quad and Kuhn [33] proposed a nested genetic algorithm to minimize setup costs and the mean flow time in batch scheduling flexible flow lines. The nested GA consists of two parts. First, an outer GA determines a target number of setups from which batch sizes are derived. Secondly, the inner GA determines a schedule for all production stages based on the outer GA parameters. The proposed GA focused exclusively on batch scheduling in a flexible flow line. Tavakkoli-Moghadam et al. [34] proposed a mimetic algorithm (MA) with a novel local search engine for solving a flexible flow line with process blocking. Computational results from the proposed MA on large-sized problems were found to be more effective when compared to classic GA.

In this paper, the flexible flow line with an objective of minimizing the makespan will be explained and formulated as an integer programming (IP) model. Due to the difficulty of obtaining the optimal solutions for medium and large-size problems, we use the Fuzzy Cell Genetic Algorithm for solving such problems. To the authors knowledge, no specific work relating to the general form of the flexible flow line scheduling problem using GA optimization or any other optimization method is published in literature to compare to the performance of the proposed method. Also, this benchmark is the first time a clear explanation is offered for generating problems with known solutions.

II. FLEXIBLE FLOW LINE SCHEDULING PROBLEM DESCRIPTION

The mathematical formulation for the two-stage flexible flow shop with the objective of minimizing the makespan was presented by Guirchoun et al. [35]. Kurz and Askin [9] proposed the mathematical model for the flexible flow line that also considers sequence-dependent setup times.

The integer programming model formulation which addresses the problem in this paper will be shown here. Let $t$ be the number of tasks to be schedule and $k$ the number of parallel machines at station $s$. $N$, $W$ and $P$ represent sets of jobs, stages and parallel machines respectively. The problem definition is presented as:

Indices

$t$: index of job to be scheduled ($t = 1, 2, \ldots , n$)
$s$: index of stations ($s = 1, 2, \ldots , m$)
$k$: index of machine ($k = 1, 2, \ldots , p$)

Parameters

\[ P_{t,s} \] = processing time of job $s$ in station $t$

\[ C_{s,k} \] = completion time of all jobs in station $s$ and machine $k$

\[ C_{mn,k} \] = completion time of all jobs in the last station $mn$ and machine $k$

\[ F_{t,s,k} \] = finishing time of task $t$ in machine $k$ in station $s$

Decision Variables

\[ x_{0,t,k} \] = 1 if job $t$ performs as first job in parallel machine $k$ and 0 otherwise

\[ x_{mn,k} \] = 1 if job $t$ performs as last job in parallel machine $k$ and 0 otherwise

\[ x_{i,j,k} \] = 1 if job $i$ performs before job $j$ in parallel machine $k$ and 0 otherwise

minimize $z$ \hspace{1cm} (1)

s.t. \[ \sum_{t \in N} x_{0,t,k} = 1, \forall k \in P \] \hspace{1cm} (2)

\[ \sum_{k \in P} \sum_{j \in N} x_{t,j,k} = 1, \forall t, j \in N | t \neq j \] \hspace{1cm} (3)

\[ x_{0,t,k} + \sum_{i \in N} x_{i,t,k} = x_{n,t,k} + \sum_{j \in N} x_{t,j,k}, \forall k \in P, \forall i, j, t \in N | i, j \neq t \] \hspace{1cm} (4)
there are machines in that stage. Eq.(1) represents the objective function which is to minimize the makespan \( z \). Eq.(2) defines the constraint to assign at least one job as a starting job to each machine in parallel for each station. Constraint (3) ensures that job \( t \) is either the first job or should be done immediately after job \( j \) in machine \( k \). Eq.(4) represents all the eligibility of each job in each machine. Constraint set (5) ensures that the completion time of all the jobs in machine \( k \) and station \( s \) should be greater than or equal to the processing time of each individual task in the same machine and station. In Eq.(6), finishing time of the first job assigned to machine \( k \) at the first station should be greater than or equal to the processing time of the first job at the same machine. Constraint set (7) enforces job \( t \) sequencing through each station after another. Eq.(8) calculates the accumulation of processing time for the set of tasks in each machine at the first station. Constraint set (9) and (10) calculate the finishing time of job \( t \) according to the equations (16) and (17).

If \( t \) is the first job in parallel \( k \) at station \( s \):

\[
F_{t,s,k} = F_{t,s-1,k} + Pt_{t,s}, \forall s \in W, \forall k \in P, \forall t \in N
\] (16)

else if \( t \) is the successor of the task \( j \) in parallel \( l \) at station \( s \):

\[
F_{t,s,k} = \max(F_{t,s-1,k}, F_{t,s-1,k}) + Pt_{t,s}, \forall t, j \in T | t \neq j, \forall s \in W | s > 1, \forall k \in P
\] (17)

Constraint set (11) links the decision variable \( z \) and \( C_{mn,l} \) to minimize the completion time of the maximum machine in parallel at the last station \( mn \). The last sets of equations (12), (13), (14), and (15) are defining the binary decision variables for the predecessor and successor job relations, first job, last job and zero for the case when the predecessor and successor have the same syntax in the problem.

III. FUZZY CELL GENETIC ALGORITHM (FCGA)

The Fuzzy Cell Genetic Algorithm (FCGA) mimics the main concepts of Meiosis and Mitosis found in human cell division. During the mitosis process, a cell with 2n number of chromosomes begins by growing (replicating) and the individual chromosomes start to align. Then, chromatids (one of two identical chromatid strands) move toward the left or right side of the cell without any physical connection. Finally, the content of the cell is divided into two new daughter cells, and the chromosomes in each are replicated to two sister chromatids. The final product of mitosis is two new cells, containing 2n chromosomes.

In a different manner, Meiosis consists of two sub-steps in the process namely meiosis I and meiosis II. The final product of this process with four cells where each one containing two distinct chromatids. Crossover happens during Meiosis I. The final step in meiosis occurs when four cells with each one contains two distinct chromatids. Mitosis and meiosis are two distinct processes in cell division for making sexual and asexual chromosomes. All these procedures are replicated the process in human cells.

The FCGA starts with random populations (Population1 and Population2) of job sequences and the order in which they will process in each stage. Chromosomes in each population are divided into two sub-populations for the mitosis and meiosis procedures using the fuzzy logic controller. Fuzzy logic intelligently controls the number of Asexual and Sexual chromosomes contained in Subpopulation1 and Subpopulation2. Individuals sent to the meiosis sub-population are going through the crossover and mutation procedures while the chromosomes sent to mitosis are going through duplication and have a smaller probability of mutation. New chromosomes in each subpopulation will be evaluated using a fitness function which calculates the makespan corresponding to that particular job sequence.

There are three metrics to assess FCGA performance that are used for controlling and prevent it from premature convergence. \( F_{\text{average}} \) defines the mean fitness distribution of the population and \( F_{\text{best}} \) (best fitness) measures the best individual (lowest fitness value for minimization problems). The worst fitness (\( F_{\text{worst}} \)) is defined as the worst individual fitness value in the population. To obtain the changing rate (\( F_{\text{change}} \)) in a population, \( F_{\text{worst}} - F_{\text{best}} \) is calculated in each iteration. It should be mentioned that the fitness value is the makespan...
for each sequence inside the chromosome. The three metrics are used as inputs for the fuzzy controller. The controller module monitors the performance of the meiosis process. If the diversity rate inside the populations decreases, the controller will start to increase the size of Subpopulation1 for mitosis. This increases solution diversity and moves away from premature convergence. However, the increased changing rate will decrease the size of mitosis candidates in Subpopulation2 to protect form diverging away from the optimal solution.

The requirement for generating data is \( n \) (number of jobs), \( s \) (number of stages), \( m \) (number of machines in parallel in each stage), the range of processing time for every \( i \)th job in each \( j \)th stage defined by \( P_t \) and \( T \). \( T_{j,k} \) is the sets of job in \( j \)th stage and \( k \)th parallel machine. It assumes the numbers of machines in each stage are the same and has to be more than one. The main purpose is to provide the sequence of jobs with no gap between the processes. The pseudocode of the problem generator is shown in Table I.

### TABLE I

**PSEUDOCODE OF THE PROBLEM GENERATOR FOR FLEXIBLE FLOW LINE DATA SETS**

<table>
<thead>
<tr>
<th>Steps</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Initialization</td>
</tr>
<tr>
<td>2</td>
<td>Input ( n, s, m, L, U )</td>
</tr>
<tr>
<td>3</td>
<td>Makespan = 0</td>
</tr>
<tr>
<td>4</td>
<td>Generate sets with ( n ) number of jobs ( T = T_{1,1}, T_{1,2}, ..., T_{s,k} )</td>
</tr>
<tr>
<td>5</td>
<td>For each machine in parallel ( k = 1, 2, ..., l )</td>
</tr>
<tr>
<td>6</td>
<td>( M_{span_k} \rightarrow P_t_{i,j} )</td>
</tr>
<tr>
<td>7</td>
<td>For each job in stage ( j ); ( j \in T_{j,k} ); ( i &gt; 1 )</td>
</tr>
<tr>
<td>8</td>
<td>( P_t_{i,j} \rightarrow \text{random} ) ( [L, U] )</td>
</tr>
<tr>
<td>9</td>
<td>( M_{span_k} \rightarrow M_{span_k} + P_t_{i,j} )</td>
</tr>
<tr>
<td>10</td>
<td>( P_t_{i-1,j+1} \rightarrow P_t_{i,j} )</td>
</tr>
<tr>
<td>11</td>
<td>End For</td>
</tr>
<tr>
<td>12</td>
<td>End For</td>
</tr>
<tr>
<td>13</td>
<td>( \text{Set}\ M_{span} \rightarrow \text{arg max}<em>{k=1,2,...,l}({M</em>{span_k}}) )</td>
</tr>
<tr>
<td>14</td>
<td>( \text{Save ( PT_{i,j} ) for each sequence; } M_{span} )</td>
</tr>
<tr>
<td>15</td>
<td>END</td>
</tr>
</tbody>
</table>

Figure 3 shows an example of 4 jobs, 2 stages and 2 machines in each stage. As it shown in the example, the final output for the problem generator is \( PT_{4 \times 2} = \begin{bmatrix} 4 & 6 \\ 7 & 6 \\ 6 & 4 \\ 8 & 4 \end{bmatrix} \) with a solution of 18 sec.

The problem generator provides solutions with no gap between jobs in each stage and at every machine. It should be mentioned that the generated solution is not an optimal solution. However, the solution provided is sufficient to compare with other procedures solutions. While algorithms have difficulty solving larger problems, the problem generator can easily generate solutions for any size problem.

The processing time for each job is randomly generated as an integer in the range of \([1, 10]\). There are 3 problem sets generated. The first set considers the number of jobs \( n \) = 5, 8, 10, 20, 30, 40, 50, 70, 100 and the number of stages \( s \) = 2, 5, 15 by assuming two machines in each stage. Secondly, a set of jobs \( n \) and a set of stages \( s \) = 20, 30, 40, 50, 100 are selected with 5 machines in each stage. Lastly, the number of jobs \( n \) = 20, 40, 100 and the number of stages \( s \) =
10, 20, 40, 60, 100 with 10 machines in each stage. Table II represents the parameters for the data sets.

### Table II

<table>
<thead>
<tr>
<th>Problem Sets Considered</th>
<th>Machine 2</th>
<th>Machine 5</th>
<th>Machine 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>#Jobs</td>
<td>5, 10, 20, 30, 40, 50, 70, 100</td>
<td>20, 30, 40, 50, 100</td>
<td>20, 40, 60, 100</td>
</tr>
<tr>
<td>#Stage</td>
<td>2, 5, 15</td>
<td>20, 30, 40, 50, 100</td>
<td>10, 20, 40, 60, 100</td>
</tr>
<tr>
<td>Processing time</td>
<td>Random[1, 10]</td>
<td>Random[1, 10]</td>
<td>Random[1, 10]</td>
</tr>
<tr>
<td>#Operation ((n \times s))</td>
<td>10-300</td>
<td>40-10000</td>
<td>40-10000</td>
</tr>
<tr>
<td>#Instance per set</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>#Instance total</td>
<td>100</td>
<td>50</td>
<td>60</td>
</tr>
</tbody>
</table>

#### V. Results and Discussion

The IP models have been solved using Cplex 12 on an Intel Xeon E5-26650, RAM 64 GB of memory. The first stopping criterion is to reach to 0.05 for the relative optimality criteria (Optcr), where Optcr defines as:

\[
\frac{|\text{Obj} - C_{\text{best}}|}{1.0 - 10 + |C_{\text{best}}|} < \text{Optcr}
\]  

In formulation (18) Obj is the current objective function value and \(C_{\text{best}}\) is the best proposed integer solution. The second stopping criterion is runtime, where CPU time is limited to 24 hours.

The IP model could only solve the following problem configurations \([m(\text{jobs}) \times s(\text{stages})] : [5 \times 2], [8 \times 5], \text{and} [10 \times 5]\), each featuring 2 machines in parallel at each stage. FCGA also found optimal solutions for this subset of problems, excluding two variants of \([10 \times 5]\) which have been solved by GAMS[38]. Figure 4 illustrates the results of FCGA and the IP model across the aforementioned subset of problems.

The remaining 70 problems with two machines in parallel that were unable to be solved by the IP model are plotted separately in Figure 5. FCGA produced better solutions than the problem generator in 59 of the 70 problem configurations. In 2 of the other 11 instances, FCGA found the same solution provided by the problem generator. The FCGA has shown outstanding search performance in a solution space for small size problems. However, the problem generator hasnt shown great performance for problem sets when compared to the IP model. Therefore, the problem generator solutions do not provide a robust reference for FCGA on larger problem instances.

In the second problem set featuring 5 machines in parallel, FCGA outperformed the problem generator in 16 of the 50 problem instances (Figure 6). FCGA solutions are measured relative to the best known solution via a loss function where loss is calculated as:  \((\text{makespan} - C_{\text{best}})/C_{\text{best}}\). Loss is accumulated for each subset of problems to assess FCGA performance.

Accumulated loss for problem instances \([30 \times 30], [40 \times 40]\) and \([50 \times 50]\) are 0.27, 0.52, and 0.42 respectively. The highest total loss value for FCGA is 0.86 for \([100 \times 100]\) problem instances. On a per problem basis this represents a small amount of loss. The problem generator solutions have shown to be an efficient reference point for large size problems.

The last part is concentrated on more complex problems featuring 10 machines in each stage which is typically larger
than problems found in the literature. As it shown in Figure 7, FCGA performs much better for the problems with the size of 
\([20 \times 20], [40 \times 40], [100 \times 10], [100 \times 20]\) and the highest total loss is less than 0.2% for all problems. However, the loss value increased by 0.77 for the problem set with size of \([100 \times 60]\). For really large size problems \([100 \times 100]\) the loss goes up to 0.77 which is still small value for the FCGA performance on that complexity.

The problem generator could provide the superior solutions for the truly large size problem and provide robust references to assess the FCGA performance.

VI. CONCLUSION

The flexible flow line problem is an important representation of many real world applications. Because of the complexity of this problem, not only it is difficult to find optimal solutions, but benchmark problems are also sparse. In this research a formulation and a FCGA procedure for solving this problem has been presented and discussed. A process for generating problem set is also developed. The performance of FCGA is tested, and the results show that FCGA can be used for solving real world size problem with satisfactory solutions.

REFERENCES


