A Multi-Level GA Search with Application to the Resource-Constrained Re-Entrant Flow Shop Scheduling Problem

Danping Lin, C.K.M. Lee

Abstract—Re-entrant scheduling is an important search problem with many constraints in the flow shop. In the literature, a number of approaches have been investigated from exact methods to meta-heuristics. This paper presents a genetic algorithm that encodes the problem as multi-level chromosomes to reflect the dependent relationship of the re-entrant possibility and resource consumption. The novel encoding way conserves the intact information of the data and fastens the convergence to the near optimal solutions. To test the effectiveness of the method, it has been applied to the resource-constrained re-entrant flow shop scheduling problem. Computational results show that the proposed GA performs better than the simulated annealing algorithm in the measure of the makespan.

Keywords—Resource-constrained, re-entrant, genetic algorithm (GA), multi-level encoding

I. INTRODUCTION

RE-ENTRANT flow shop scheduling problem (RFSP) is an important industrial problem that faced by many semiconductor plants, repairing companies or electronic companies. In such assembly lines, the work in process will repeatedly pass through some of the work stations at different stages of operation with the same order. When each machine processes the set of all jobs in the same order too, this problem evolves into the re-entrant permutation flow shop scheduling problem (RPFSP), which is the special case of RFSP. As RFSP is known to be NP-hard, when resource constraints are taken into consideration, the searching space is highly constrained. In some production lines, the production execution mode has direct relation with the re-entrant chance. Take the electro-coating as example, high-end machines such as high pressure sprayer offer high performance and consume more paint and are less prone to rework, whereas skilled worker can make up the resource consumption with more time allocation but with a higher possibility of re-work. However, the welding process is the opposite case. Manual work promises better welding within the tolerance of quality standard, whereas automatic welding costs more filler material and has higher chance of rework. Hence, according to the application, the practical importance and theoretical challenge, there is still a strong motivation to explore the scheduling problem of alternative execution mode which is related to the rework probability. This kind of scheduling solution has a higher pay-off when there is limited expensive equipment and high skilled labor. In general, meta-heuristic methods differ from exact methods in the sense that not only the meta-heuristics methods can obtain a nearly optimal solution within a reasonable computational time but also it can manipulate some solutions to attain solutions with higher quality.

Therefore, the resource-constrained re-entrant scheduling problem in the flow shop needs such tools. Introduced in [1], genetic algorithm serves as an important meta-heuristic strategy to solve the hard optimization problems. So in this paper a genetic algorithm with multi-level encoding approach is proposed to solve the resource-constrained re-entrant flow shop scheduling problem where the job’s re-entrant possibility depend on the execution mode and resource consumption. The proposed GA will integrate the tasks of job scheduling and execution mode selection, which are subject to the precedence constraints and capacity constraints of machine and human resource, such that the makespan of the schedule is minimized.

The rest of this paper is devoted to describing the problem’s characteristic with the multi-level encoding and its application to the practical problem. In section II, the related work is discussed, and in section III the proposed method is presented. Section IV conducts the proposed method with the case study and the results are presented and analyzed. Section V sums up the study and remarks the future directions.

II. LITERATURE REVIEW

A. Related Work of Re-entrant Flow Shop Scheduling

The problem of re-entrant flow shop scheduling can be traced to the work of Graves et al. [2] who is the first one to systematically define the RFSP term as the problems happened in the shop area with a sequence of operations are repeatedly processed by machines.

However this problem was only studied broadly since 2000 with the demanding manufacturing requirement from the semiconductor industries or integrated circuit fabrication or thin film transistor liquid crystal display process. Exact method, heuristic and meta-heuristic methods all have been proposed for solving the re-entrant flow shop scheduling problem. However, the comprehensive analysis of this problem is not done till the review paper of Lin and Lee [3] which classifies the re-entrant problems and analyzes the optimization methods of re-entrant manufacturing scheduling.

Mostly exact methods formulate the re-entrant flow shop scheduling problem as the integer programming problem and simulate the efficiency of the new proposed approaches. However, the computational complexity is high as long run time is used to narrow down the space within the constraints. The referenced work can be found from [4-15].

Heuristic techniques such as dispatching rules and constructive heuristics have also been applied to the re-entrant flow shop scheduling problem. Recent work that applied the dispatching rules in the semiconductor industries are found from [16-18]. They always tested the proposed modes in terms of simulation. Though these heuristic methods produce good quality solutions, they are not extensively used during the period of 2000-2007 as the running time are often large and increase rapidly with the size of the problem increases.

Danping Lin is with Nanyang technological university, Singapore e-mail : Lind0020@e.ntu.edu.sg
Meta-heuristic, especially genetic algorithm (GA), evolved to be the new trend for the re-entrant flow shop scheduling problem. GA gets the edges over other techniques in the area of large-scale complexity like the very large scale integrated (VLSI) circuit manufacturing or repairing industries which are highly re-entrant, and involve hundreds of machines, restrictions and processing steps. Therefore, [19-23] used GA to solve the re-entrant problem and mainly focus their objectives on minimizing the makespan. Another meta-heuristic method tabu search has also been used, which can be found from [24-25].

Going through the related work, we found that exact methods and meta-heuristics are the main tools based on the frequency of used times and most of them adopted minimizing the makespan as their objectives. However, most of the works assumed unlimited buffer or no resource constraint in the processing procedure, or some of the work relaxed these constraints and decomposed the problem into sub-problems; these approaches isolated the resource constraint from the re-entrant flow shop scheduling. Only the work of Scholz-Reiter et al. [15] that fully takes the resource constraints into account. In their work, machines and labor resources are taken into consideration, priority rules and rule combinations are simulated both for small problem instances and long term system behavior with a time span of ten years. The results revealed that the performance of priority rules depends on the utilization level of the system, thus a proper control system is crucial to select and combine the priority rules.

B. Related Work of Flow Shop with Constrained Resource

Generally the works of resource-constrained flow shop can be classified into two types in terms of objectives. The first type related to allocating and scheduling the limited resource to the flow shop. Daniels et al. [25] discussed the feasibility and benefits of dynamic assignment of partially flexible labor resources, linked skill matrices with self-buffering production line and eventually proved the benefit of complete flexibility can be achieved by a reasonable small amount of carefully selected partial flexibility. Similar work is done by Neubert and Savino [26].

The second type of flow shop with limited resource mainly concerns the jobs scheduling problem, while the resource constraints are only one of their considerations. The example can be found from Janiak [27], whose main objective is to minimize the job overall completion time. In his model, the duration of each operation on certain machines is a linear function of the allocated part of a constrained resource. The proposed algorithm is based on disjunctive graph theory and brand & bound algorithm. However, resource allocation is regarded as one of the concerns for scheduling. The similar work has been done by Cheng and Janiak [28] and Leu and Hwang [29]. Although there are a lot of research has been done on flow shop scheduling problem, according to authors’ knowledge, no studies have done connecting the resource constraint with the re-entrant possibility but it is quite common for the flow shop circumstance. Therefore, we will propose a GA-based technique that addresses the re-entrant flow shop scheduling problem with resource constraints. Our GA is designed to perform the combination work of job scheduling, resource satisfaction and re-entrant possibility control. Our approach will offer a novel way of chromosome encoding for this specific problem.

III. MULTI-LEVEL ENCODING GENETIC ALGORITHM

A. Multi-level Chromosome Encoding

There is a strong interdependence between the execution mode selection and the re-work possibility, and it is not clear in which order they should be performed. Therefore, it is inappropriate to separate the schedules with execution mode selection and re-entrant possibility control. Instead they should be considered concurrently as the search space for the combined task is large and discrete. GA is known to work well on such complicated problem and the inherent heuristic character of GA allows the synchronous exploration at the same optimization iteration. That is the main reason we choose GA. Moreover, when the integrated optimization requirement reflecting in the GA’s chromosome encoding, information need to be kept intact and undistorted as well as easy to inherit. Compared with one-level chromosome encoding, when the number of the re-entrant job is dynamic, the subsequent length of the chromosome is unfixed. Therefore, there can be information mixing among different types of data when carrying out the crossover or mutation operation. To the opposite, the multi-level encoding reserves the data into different levels according to the data types. So the length of each level is allowed to be different. Furthermore, crossover or mutation is implemented level by level, which essentially stops the chance of data mixing and distortion. At the same time, efficient repair operator is used to revise any infeasible solution. The detail of the information representation is shown in Fig. 1.
meaning, level 3 and 4 are the representation of the machine and labor consumption under two different execution modes. During the encoding, GA will randomly choose either mode 0 or mode 1 as its execution mode, thus its re-entrant possibility and corresponding resource consumption will be tracked. Assumed the combination of execution modes in Fig. 1(b) fulfills the resource constraints, it is obvious that the job 1, 3, 5 have the rework possibility more than 50 percents, which means all these jobs need to be reworked. The re-entrant jobs will be given other indexes to differentiate with the original job indexes. Of course, no matter for the original jobs or those re-entrant jobs, they should within the overall resource constraints. Therefore, there are actually nine jobs need to be scheduled with the aim of minimizing the total makespan. After that GA can do the crossover and mutation again just for the level of job sequences. It is worth to note that the crossover and mutation operators are executed level by level.

In this meaning, GA will randomly generate an initial population for each generation. During each generation, two parents are chosen to propagate new offspring by applying the crossover and mutation. Later the propagation effect is evaluated by the fitness function. At the same time the function crossover and mutation operators are executed level by level.

In order to show the effectiveness of the proposed method, the experiment is based on real case and compared with the homogeneous algorithm of simulated annealing algorithm.

EG (Alias) Ltd is one of the magnates of providing various electronic technology and services in broad domain. In south-east Asia, it endeavors in offering a wide range of repairing service to huge energy parts. Their repairing workflow is the typical flow shop as the parts will go through receiving inspection, disassembly, repairing procedure and final assembly following the same order. Before parts moving on to the new workstations, they will be inspected. When new faults are found they need to return to the previous workstations to be reworked. The rework possibility is not constant, because
the execution modes of the previous workstations will influence the consumption of machine time and worker time, and the re-entrant chance. Take the power nozzle repairing procedure as example, if the job’s due date is not so tardy, the workstation may choose the more complicated repairing mode that will go through 40 steps. Under this mode, a typical piece of power nozzle will cost 400 units of worker time, 120 units of machine time with 9 percent of rework possibility. If the job is emergent, the same piece of part can be done applying the compact mode that only goes through 13 steps, consuming 215 units of worker time, 420 units of machine time with as high as 15 percent of chance to be reworked. Whatever the original jobs or the re-entrant jobs, the differences of the execution modes for them are the processing time in some particular stages and the consumption of the machine and manpower resources and some of the parts do not need to pass all the machines.

So the proposed multi-level GA is simulated to minimize the makespan for this plant. The experiments are conducted into 20-, 30-, and 40-job sizes according to the plant’s current job classification. The experiments are implemented in the MATLAB 7.0 on a Pentium 4 3.17GHz PC with 3.25G MB memory. The parameter setting is tuned based on the foregone experiments. The size of the GA population was set as 200, the two-point crossover with crossover probability is set as 0.8, and the Gaussian mutation with mutation probability is set as 0.4. Each of the GA runs is stopped after 100 generations. While for the SA, 10 runs are conducted with fast annealing function and stopped after 200 generations. On the other side, based on the real data’s distribution, the processing time in the six workstation may choose the more complicated repairing mode as example, if the job’s due date is not so tardy, the job will go through 40 steps. Under this mode, a typical piece of part will cost 400 units of worker time, 120 units of machine time with 9 percent of rework possibility. If the job is emergent, the same piece of part can be done applying the compact mode that only goes through 13 steps, consuming 215 units of worker time, 420 units of machine time with as high as 15 percent of chance to be reworked. Whatever the original jobs or the re-entrant jobs, the differences of the execution modes for them are the processing time in some particular stages and the consumption of the machine and manpower resources and some of the parts do not need to pass all the machines.

In the simulation tests, we give the mean, standard deviation and the best performance referring different problem size respectively. Table I shows the solutions found by GA and SA in these ten runs under the resource and re-entrant constraints. Ideally, the GA-based method should converge to the optimal or near-optimal solution in each run. Computing the average of the best solutions found over the ten independent runs provides an indication of the robustness of the search strategy. In ten out of the twelve tests, GA works better than the SA. For the two out of twelve tests showing GA is inferior to SA, the differences of the fitness between them are less than 0.3%. This illustrates that the proposed GA approach is able to consistently converge to high-quality solutions in every GA run for different problem sizes.

V. CONCLUSION

This paper presented a multi-level encoding genetic algorithm for re-entrant flow shop scheduling problem with resource constraints. In the proposed approach, the multi-level encoding translates the data into different levels according to data types and executes crossover and mutation individually for each level. The proposed method has demonstrated better performance with regard to the makespan when compared with simulated annealing algorithm.

The superiority of the multi-level chromosome encoding lies in:

1) It provides integrity and flexibility to perform the combination of searching task.
2) Multi-level encoding enables user to store more information about the problem and this encoding is more clear and visualized.
3) As the re-entrant job number is dynamic, different levels have their own operators. On one hand, it helps to keep the information independently and integrated. On the other hand, it expedites the chromosome converge to the better optimal as one group of multi-level chromosome represents one solution, it does not need to decode.

Referring to the result differences of GA and SA, they may come from the inherited characteristic of the algorithms themselves. SA traverses the search space by applying local transformations to the solution representation, while GA’s search ability is decided by the chromosome representation and search operators. In the first place, the multi-level encoding is problem-specific encoding, which stores the information according to the data types so that the crossover and mutation operators can be implemented level by level without distorting and mixing different types of data. Second, when decoding the chromosomes, it is vivid and easy as one the one group of multi-level chromosomes directly represents one solution.

Future research may include applying the method to larger problem sizes and broadening the application into semiconductor industries. Improvements to the method could be explored, such as the parameters tuning or the sensitive analysis.
REFERENCES


