Technical Paper

Heuristic algorithms to minimize total weighted tardiness with stochastic rework and reprocessing times

Venkatesh Arasanipalai Raghavan [a], Sang Won Yoon [b,*], Krishnaswami Srihari [b]

[a] Optimization Engineer, Philips Lumileds, 370 W Trimble Road, San Jose, CA 95131, United States
[b] Department of Systems Science and Industrial Engineering State University of New York at Binghamton, Binghamton, NY 13902, United States

ARTICLE INFO

ABSTRACT

This paper introduces a mathematical model for the scheduling problem with stochastic rework and reprocessing time, which is typical for electronics manufacturing services (EMS) providers. Since rework and reprocessing of jobs take more time resulting in missing the due dates, this research aims to determine the optimal job sequence on the machines. A heuristic methodology is developed that takes into consideration the *Total Estimated Processing Time* (TEPT), a linear combination of processing, rework, and reprocessing times. Jobs with different configurations of processing, rework, and reprocessing times representing a High-Mix–Low-Volume (HMLV) setup are tested on a single-machine job shop system to understand the effectiveness of the shortest TEPT algorithm (STEPT). Improvements to the STEPT algorithm are made and the modified STEPT algorithm (MSTEPT) is tested using a single-machine job shop setup with a larger number of jobs. Experimental results indicate that the proposed algorithm outperforms different commonly used dispatch rules in terms of solution quality and computation time. Then, experiments conducted on a multi-machine job shop setup with larger number of jobs also indicate the superior performance of the MSTEPT algorithm.

© 2014 The Society of Manufacturing Engineers. Published by Elsevier Ltd. All rights reserved.

1. Introduction

The electronics manufacturing process is often faced with different types of job failures, which could occur during printed circuit board (PCB) assembly and final product assembly processes. For instance, defects at the PCB level include solder bridging, voiding, missing components or wrong orientation and electrical shorts or opens, and those at the final product assembly include improper alignment and connection of PCBs, mechanical defects and other quality related issues [1]. Most of the defects are captured during either different testing processes or visual inspection. When assembly defects are identified, the products are then reworked to restore them to original operating conditions. Rework could be performed either on a different rework machine, referred to as *off-line* rework, or on the same machine, referred to as *in-line* rework. For jobs undergoing off-line rework, they would have to return to the same manufacturing operation to undergo reprocessing, which could require retesting or visual inspection. For jobs undergoing *in-line* rework, both rework and reprocessing occur at the same machine (or manufacturing operation).

When the products return for rework or reprocessing, there would be other products either being processed at the same operation or waiting in the queue to be processed. Moreover, to process the returning products, an additional set-up time could be involved, such as setting up the right pick-and-place program or the right reflow profile recipe. These cause the failed units to spend more time in the production line. EMS providers generally deliver services to multiple customers for various product types, such that the large number of product types gives rise to more opportunities for defects, which results in the increased job tardiness. In most cases, the jobs are processed based on their due dates, job priority or weights, or some other criteria such as FIFO (First In First Out). If special attention is not given to those jobs that have a higher chance of failure, job tardiness would be increased. Even though there has been considerable research on the scheduling problem, there is limited research pertaining to scheduling in electronics manufacturing with stochastic rework and reprocessing times.

Therefore, the objective of this research is to develop an integer programming model to minimize the total weighted tardiness, considering rework and reprocessing times based on their probabilities. The total time that includes the processing time, as well as the rework and reprocessing times, is referred to in this research as the Total Estimated Processing Time (TEPT). Since minimizing the total weighted tardiness is NP-hard even in a single-machine

* Corresponding author. Tel.: +1 6077775935.
E-mail address: yoons@binghamton.edu (S.W. Yoon).

problem, the *Shortest Total Estimated Processing Time* (STEPT) algorithm is developed in this research. The proposed algorithm have been experimented with High-Mix–Low-Volume (HMLV) job shop setups to evaluate the effectiveness of the algorithm. Then, the STEPT algorithm has been improved and updated to deal with a large number of jobs, which has also been tested. There are several assumptions made in this research. It is assumed that both in-line and off-line rework probabilities and reprocessing time distribution for all the jobs are available. Typically, the EMS manufacturing system keeps track of rework and reprocess rates for every unit. Also, it is assumed that failure and rework of any job at any operation occurs only once at the maximum and a failed job returns to original operating conditions if the rework has been performed.

The remainder of this paper is summarized as follows. A brief review of literature is presented in Section 2. The integer programming model to minimize the total weighted tardiness is explained in Section 3. The STEPT algorithm and the modified STEPT algorithm are developed in Section 4. Experimental results and analyses are discussed in Section 5. Finally, conclusions, limitations, and future work are outlined in Section 6.

2. Literature review

The scheduling problem is complex because the number of solutions could grow exponentially with the problem size, making it computationally intensive [2,3]. A number of researchers have discussed the computational complexity of scheduling [4–6]. This complex nature of the problem has attracted the interest of researchers, which has led to a number of research articles in this area. As a result, many researchers have developed different methods to obtain optimal to near-optimal solutions for scheduling. One of the earliest methods was the use of priority rules, which have had an influence on the scheduling process [7]. It has been shown that the Shortest Processing Time (SPT) rule dominated several other rules and was considered as the standard rule for comparison [8].

This research focuses on scheduling jobs in an electronics manufacturing system. From the literature, one research aimed at scheduling Printed Circuit Boards (PCBs) in a component interchange machine, such that the total number of component switches was a minimum [9]. Studies to optimize the feeder set-up policy and allocation of feeders to slots in pick-and-place machines have also been conducted [10,11]. There were also some studies that have considered batch processing test operations in the electronics manufacturing facility and aimed to minimize the makespan [12–14].

In order to solve these problems, several metaheuristic approaches have been studied; e.g., Genetic Algorithm (GA) and Ant Colony Optimization (ACO) at Environmental Stress Screening (ESS) chambers used in PCB manufacturing [15]. Different research on two batch processing machines in a flow shop has been conducted to minimize the makespan using Simulated Annealing (SA) algorithm [16]. It can be observed that several studies in the electronics manufacturing domain have been conducted to minimize the makespan. There has not been much research with an objective of minimizing the total weighted tardiness (TWT) in the electronics manufacturing domain. Companies can develop a good will with their customers by adhering to the due dates, and one of the ways to ensure due date conformance is to monitor the weighted tardiness as a performance measure [17]. However, it has been seen that the TWT is one of the most complex objective functions to minimize in the job shop scheduling problem (JSP) [18]. According to the literature, the single machine TWT minimization problem itself is NP-hard [19]. As a result, several studies have shown how to minimize the TWT using heuristic algorithms and compared their performance against dispatch rules such as Earliest Due Date

(EDD), Weighted EDD, SPT, and Weighted SPT [20–23]. A similar approach will be followed in this research also for comparing the performance of the proposed algorithm.

The research conducted over recent years can be classified as a problem- or methodology-based research. Although several methodologies have been developed to address different areas in the scheduling research, there has been little research that has addressed jobs that fail and undergo rework. Dispatch rules have been developed for a photolithographic area with on-line rework strategy to balance system load and improve throughput [24]. However, their rework time will be equal to the time to rework a wafer at the resist stripping (rework) process and will not vary with jobs. The objective of another research article that considered rework was to determine the sequence of batch sizes to minimize the makespan of a single product [25]. Their model considered a single product, and hence the rework time was assumed to be identical for all units. Additionally, their model was a two-stage model with the second stage being the rework stage. The weighted tardiness and makespan were minimized in a flow shop with re-entrant jobs [26]. However, they neither considered the effects of rework time or reprocessing time, nor developed any dispatching rules or algorithms to efficiently handle re-entrant jobs. In addition, a parallel machine scheduling problem has adapted rework and reprocessing times [27,28]. Their research aimed at scheduling jobs to parallel and available machines to reduce their chances for failure. However, they have not developed any mathematical model and considered any preemptive rule in their model. Since there were multiple machines to process, the issues arising when having only one machine remains unaddressed.

Also, jobs undergoing rework in a branched flow-shop have been modeled to reduce tardiness [29]. They identified time buffers in the second line to compensate for excessive wait time due to rework of jobs from first line. A simulation-based real-time dynamic scheduling mechanism has been used for the semiconductor industry [30]. The best combination of rework strategy and dispatching rule was determined. In summary, there has been limited research focusing on the electronics manufacturing industry with jobs undergoing rework and reprocessing.

3. Mathematical model

Generally, the cycle time of jobs increases due to rework and reprocessing. This could delay on-time job completion. This research presents a mathematical model using the TEPT, which will be addressed in this section. A generic flow of jobs between machines (or operations) is given in Fig. 1. All operations at which jobs could fail will have to go through the in-line and off-line rework loops as shown in Fig. 1. A list of nomenclatures is summarized in Table 1. The time index t varies from 0 to T, where T is the finish time of the last job at its last operation, FT_{nm}.

3.1. Rework and reprocessing

It is assumed that any job i could undergo in-line rework at operation j with a probability $(r_{ij} \cdot \mu_{ij})$. Therefore, off-line rework probability for the same job i is $r_{ij} \cdot (1 - \mu_{ij})$. Upon returning to operation j after off-line rework, the job would undergo reprocessing (given by the processing time P_{ij}). A job could also return to operation j for in-line rework followed by reprocessing. Thus, the different categories of jobs that could be present in queue at operation j are new jobs, jobs that underwent off-line rework and jobs waiting to undergo in-line rework.

The Expected Time at Return at operation j spent by a job i, $E[TR_{ij}]$, is the time that the job would have to spend for either