G Model
JMSY-236; No. of Pages 11

ARTICLE IN PRESS

Journal of Manufacturing Systems xxx (2013) xxx–xxx

Contents lists available at ScienceDirect

Journal of Manufacturing Systems

journal homepage: www.elsevier.com/locate/jmansys

Technical paper

Multi-objective real-time dispatching for integrated delivery in a Fab using GA based simulation optimization

Xiaokun Chang[a], Ming Dong[a,*], Dong Yang[b]

[a] Department of Operations Management, Antai College of Economics & Management, Shanghai Jiao Tong University, 535 Fahua Zhen Road, Shanghai 200052, PR China
[b] Department of Information Management, College of Management, Shanghai Donghua University, 1882 Yan'an Road, Shanghai 200051, PR China

ARTICLE INFO

Article history:
Received 11 December 2011
Received in revised form 16 June 2013
Accepted 9 July 2013
Available online xxx

Keywords:
Integrated delivery
Dispatching rule
GA-based simulation optimization methodology
Response surface methodology

ABSTRACT

In a wafer fabrication Fab, the "integrated delivery", which integrates the automated material handling system (AMHS) with processing tools to automate the material flow, is difficult to implement due to the system complexity and uncertainty. The previous dispatching studies in semiconductor manufacturing have mainly focused on the tool dispatching. Few studies have been done for analyzing combinatorial dispatching rules including lot dispatching, batch dispatching and automated guided vehicle (AGV) dispatching. To handle this problem, a GA (genetic algorithm) based simulation optimization methodology, which consists of the on-line scheduler and the off-line scheduler, is presented in this paper. The on-line scheduler is used to monitor and implement optimal combinatorial dispatching rules to the semiconductor wafer fabrication system. The off-line scheduler is employed to search for optimal combinatorial dispatching rules. In this study, the response surface methodology is adopted to optimize the GA parameters. Finally, an experimental bay of wafer fabrication Fab is constructed and numerical experiments show that the proposed approach can significantly improve the performance of the "integrated delivery system" compared with the traditional single dispatching rule approach.

1. Introduction and literature review

Semiconductor manufacturing is one of the most sophisticated manufacturing processes with characteristics of large processing steps, re-entrant material flows, and batching processing. In the past decades, the dispatching decision problems of processing tools and material handling in wafer fabrication are usually investigated separately due to the modeling complexity. However, an effective dispatching mechanism which properly integrates tools dispatching with AMHS dispatching is becoming a more widely used strategy to achieve high performances in semiconductor wafer fabrication systems.

Most approaches to the semiconductor manufacturing scheduling problem can be classified into four categories: heuristic rules, mathematical programming techniques (such as branch and bound, Lagrangian relaxation and queuing network model), neighborhood search methods (such as Tabu search, genetic algorithm and filtered beam search) and artificial intelligence techniques (such as artificial neural networks and expert/knowledge-based systems) [1].

In recent years, many studies have focused on the dispatching rules for scheduling semiconductor wafer fabrication systems. Heuristic rules and mathematical programming techniques are widely used for this problem. Yang and Chang formulate a multi-objective model for IC (Integrated Circuit) sort and test [2]. In their study, Lagrangian relaxation is used to look for an approximate Pareto boundary and a new algorithm is designed to solve the dual problem. Li et al. propose a dispatching rule to improve on-time delivery without decreasing throughput and increasing cycle time [3]. Wu et al. develop a line balance-starvation avoidance (LB-SA) algorithm based on a proposed simplification model of the processing route for a Fab with machine-dedication features [4]. Monch and Driebel also propose a heuristic rule, named modified shifting bottleneck heuristic, for wafer fabrication Fab [5]. Among these approaches, dispatching rules, by far, are the most commonly used tools for shop floor control. The reason behind this is that they are simple to implement, quick in reacting to the changes encountered on the shop floor, easy to understand and require a low computational load [6].

Pierce and Yurtsever present a value-based concept. They mainly concentrate on generating profits [7]. Based on that, Hsieh and Hou develop a production-flow-value-based job dispatching rule (PFV) by the theory of constraints (TOC) for wafer fabrication [8]. The TOC and profitability costs estimation of a WIP-wafer lot are derived to prioritized jobs based on their profitability. Although

* Corresponding author.
E-mail address: mdong@sjtu.edu.cn (M. Dong).

0278-6125/$ – see front matter © 2013 The Society of Manufacturing Engineers. Published by Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jmsy.2013.07.001

dispatching rules are usually sub-optimal and myopic, over the years, researchers have successfully introduced advanced rules that are capable of improving multiple performance measures simultaneously [6]. Dabbas and Fowler develop a composite dispatching rule for each station, which combines multiple dispatching rules, including local and global rules, into a single rule [9]. It is designed to simultaneously maximize the on-time delivery and minimize the variance of lateness and cycle time by using a linear combination with relative weights. Bahaji and Kuhl also propose a composite dispatching rule and prove it to be robust for the average and variance of flow time, as well as due-date adherence measures [10]. Another new weighted composite rule suggested by Yang et al. also take into consideration the general advantage of the earliest due date (EDD) and shortest processing time (SPT) rules, with the weights determined from the historical data [11]. However, those rules above are designed for lots in single processing stations and may not perform the best for different scenarios.

Overhead monorail systems are widely used in wafer manufacturing systems due to their efficiency in lean designs, gentle handling of wafers and efficient use of overhead space [12]. It has been demonstrated that they can significantly improve the inventory storage and material handling system reliability [13]. Christopher et al. indicate that both processing tool and vehicle dispatching rules and their interaction have a significant impact on Fab performance if the AMHS is not extremely over or under utilized. It also has been shown that the combination of dispatching rules is highly dependent on the specific Fab [14].

In order to solve the scheduling problem that integrates processing tools and vehicles, Lin et al. propose a hybrid push/pull dispatching rule, including the procedures for vehicle selecting lot (VSL) and lot selecting vehicle (LSV), to improve the photobay performance [15]. Min and Yih propose a combined simulation and neural network approach in which simulation is used to collect data related with the system status, performance and change in dispatching rules, and then its results are fed into a neural network to obtain related knowledge [16]. Tyan et al. propose a state-dependent policy for achieving better system performance [17]. However, these methods are not capable of handling complex stochastic semiconductor manufacturing environments.

Through a survey on the use of discrete event simulation for manufacturing system design and operation, Smith shows that simulation has proven to be an extremely useful analysis tool for developing dispatching rules [18]. Um et al. gives a simulation design and analysis of a flexible manufacturing system with an automated guided vehicle system (AGVs) [19]. Hung and Chen explore a simulation-based dispatching rule and a queue prediction dispatching rule for searching better dispatching rules to reduce flow times, while maintaining a high machine utilization [20]. Kim and Jeong present a simulation-based real-time scheduling methodology which considers both simulation models and decision time points for reselecting new rules [21]. Sivakumar also develops a discrete event simulation-based on-line multi-objective scheduling approach, which includes the use of a linear optimization algorithm and automatic simulation model generation, in a complex manufacturing environment [22]. Kim et al. present a simulation-based real-time scheduling (SBRTS) methodology in which lot scheduling rules and batch scheduling rules are selected from candidate rules based on information obtained from the simulation [23]. Jelong et al. present a hybrid approach that combined GA with the simulation [24]. However, only single performance indicator, namely the maximum completion time for the last job, is optimized in their study.

The contribution of this paper is twofold. First, it mainly concentrates on the "combinatorial dispatching" which has rarely been studied. Different from previous study on single performance indicator or weighted composite dispatching rules, a representative Pareto optimal solution subset is explored to handle multiple objectives. A GA-simulation procedure is applied to obtain this solution subset. This methodology could provide users the opportunities to select appropriate solutions according to their preferences. Second, a GA parameters optimization process is developed to handle complex environments. This process can improve the performance results by optimizing GA parameters even with the dramatically changing environments.

The remainder of this paper is organized as follows. In Section 2, the proposed methodology is given. Related dispatching rules, GA parameters optimization by response surface methodology and simulator parameters optimization will be presented in this section. Sections 3 and 4 describe GA controller and the simulation model in details, respectively. In Section 5, numerical experiments used to show that the proposed methodology can efficiently improve the overall system performance. Finally, Section 6 concludes the paper.

2. The GA based simulation optimization methodology

The proposed methodology that integrates processing tools and the AMHS is applied to real-time multiple-objective scheduling problems in semiconductor wafer fabrication systems. It mainly consists of two parts: the on-line real-time scheduler and the off-line scheduler, as illustrated in Fig. 1. The on-line scheduler is mainly employed to monitor and convey the information between users, semiconductor wafer fabrication system and off-line scheduler. It interacts with the off-line scheduler in ways of initializing the off-line scheduler and accepting the combinatorial dispatching rules optimized by the off-line scheduler. Then it applies the accepted combinatorial dispatching rules to the real system. Additionally, it can also be used to determine the points of time when new rules are reselected.

Kim and Jeong [21] provide three types of points of time for reselecting dispatching rules: the beginning of each planning horizon, the time when a major system disturbance occurs and the time when the differences between actual and estimated performance values exceed a pre-determined limit. They also categorize the system disturbances into two levels: major disturbances and minor disturbances. Major disturbances include arrivals of urgent jobs and major machine breakdowns that require a long repair time or for which the repair time cannot be estimated. On the other hand, the machine breakdowns for which the repair time is estimated to be short are considered as minor disturbances. In this paper, this method is adopted to determine the points of time when rules are reselected.

The off-line scheduler consists of four modules: GA parameters optimization module, simulation parameters optimization module, GA controller and the simulation model. The former two modules are used to search for better GA parameters (such as population size, crossover rate, mutation rate and stopping criterion) and simulation parameters (simulation length and number of replications), respectively. They are often conducted independently of the GA controller after receiving trigger signals from users or the on-line scheduler. The triggering signals mainly come from the system environment changes (such as material plan change and processing tool breakdown) or user requests.

GA controller and the simulation model are the most important modules of the off-line scheduler. Iterative runs of GA controller and the simulation model compose the iterative GA-simulation procedure which aims at finding the optimal combinatorial dispatching rules. The procedure is as follows. First, the GA controller produces different chromosome individuals which encode specific code numbers and meanwhile the simulation model selects the corresponding dispatching rules according to the code numbers.