Proceedings of the 2003 IEEE
International Conference on Robotics & Automation
Taipei, Taiwan, September 14-19, 2003

Colored Timed Petri-Net and GA Based Approach to Modeling and Scheduling for Wafer Probe Center

Shun-Yu Lin, Li-Chen Fu, Tsung-Che Chiang, and Yi-Shiuan Shen

Department of Computer Science and Information Engineering
National Taiwan University, Taipei, Taiwan, R.O.C.

Abstract—*In this paper, we propose an architecture to simulate Wafer Probe. We use a modeling tool named CTPN (Colored-Timed Petri Nets) to model all testing flow. With CTPN model, we can predict the delivery date of any specific product under some scheduling policies efficiently and precisely. In the scheduling phase, we combine two popular methods to construct high-quality schedules. One is to select machines for lots and the other is to select lots for machines. In each method, we use the GA-based approach to search for the optimal combination of a number of heuristic rules. This CTPN-based GA scheduler and selection method enlarges the solution space and helps us to find the good solution so as to meet the requirements in the complicated environment.*

1. Introduction

In recent years, semiconductor manufacturing industry grows rapidly and it has become one of the most competitive fields. Since the wafer probe can be viewed as the final step of wafer fabrication and various kinds of products will be tested in this step, wafer probe center becomes more important.

In such a competitive market, wafer probe centers must try hard to reduce cycle time, improve quality, and meet the target. In addition, they have to respond quickly to orders and meet due dates. This paper addresses the problem of scheduling in the wafer probe center. If we want to meet requirements of customers, we have to precisely estimate the cycle times of all orders as well as other performance measures (Work In Process, target, machine utilization) first. Based on these measures, the manager or engineers in the wafer probe center can control the operation easily. In order to accomplish the task, a wafer probe center model based on CTPN (Colored Timed Petri Nets) is proposed in this paper. This model not only helps us to evaluate these performance measures, but also helps us to develop and evaluate our scheduler easily. Most of the literature that addressed scheduling issues in wafer probe center has appeared in the last decade. Uzsoy *et al.* [6] developed production scheduling algorithms for semiconductor test operations through constructing the disjunctive graph representation of the testing system. Chen [5] modeled the scheduling problem for IC sort and test facilities as an integer programming problem and used the Lagrangian relaxation technique to solve it. Among them, the modeling tool based on Petri Net is the most commonly used one, which is also adopted in this paper. Xiong *et al.* [3] proposed and evaluated two Petri-net based hybrid heuristic search strategies and their applications to semiconductor test facility scheduling. In Allam *et al.* [4], hybrid Petri nets were adopted as tools for modeling and simulation of semiconductor manufacturing systems. The colored timed Petri net (CTPN) is used to model the furnace in the IC wafer fabrication [2] and in the whole wafer fabrication manufacturing system [1].

There are many papers on planning and scheduling. In recent years, an interest in using probabilistic search methods to solve job shop problem has been growing rapidly, such as simulated annealing (SA), tabu search (TS) [11], knowledge-based approach , and genetic algorithms (GA) [13]. They also discussed chromosome representation in detail. Riyaz Sikora [11] and Lee *et al.* [8] focused on solving the scheduling problem in a flow line with variable lot sizes. Cheng *et al.* [12] have surveyed related topics on solving the job shop problem using GA. Murata *et al.* [10] examined hybridization of the genetic algorithm with other search algorithms in flow shop scheduling.

The organization of this paper is described as follows. In Section 2, the environments of wafer probe center as well as overview of a scheduler are given. Then, the wafer probe model based on CTPN (Colored Timed Petri Nets) considering some practical issues in modeling is described in Section 3. Then, modeling applications, prediction and scheduling, are described in Section 4. In section 5, the implemented system is described. Some comparisons are made to validate the proposed methods. Finally, a brief conclusion is given in Section 6.

2. Environments of Wafer Probe Center

Each wafer contains several tens or even hundreds of individual devices, named die. Because the process in wafer fabrication is complicated, the goal of the probe process is to ensure that wafer is defect-free by using automated testing equipment to examine each circuit and to determine whether it is operating at the required specification. In wafer probe center, it includes four major stages: sort, laser repair, ink, and bake. Sort stage tests the electric functions of circuits and layout

between circuits in each die, laser repair stage tries to repair the defective die and to improve the yield, ink stage marks defective die for packaging, and bake stage tries to dry the ink and to clean the surface of wafer.

There are many conditions needed to be met when facing scheduling in the probe center. The first one is capability limitation, which means that each lot must be sorted in the correct equipment. Because the capability and capacity of each equipment are not the same and each equipment may have multiple capabilities, the scheduler must make assignments of equipment and lots considering the constraints from capability. The second one is the accessory limitation. For the sort stage, there are two accessories, namely, probe cards and load boards. The accessories must work with the correct equipment type and the correct product type. The third one is the limitation resulted from test programs. Test program is used to sort each product, and they are different for different products. According to these three factors, we must change accessories and test programs with respect to the equipments and the product types. Too frequent changeover will make equipment utilization low. On the other hand, we may produce insufficient quantity of some products if we limit changeovers inappropriately. As a consequence, the scheduler must plan the equipment usage well.

3. Modeling of Semiconductor Probe Center

A wafer probe center can be considered as a job shop containing a large number of equipments. Simulation is one method of solving scheduling problem. With simulation, it must build a detailed model to simulate the flows of products. In our method of Colored timed Petri Net (CTPN), we can easily use color and time attributes in Petri Net to model the wafer probe center for meeting limitation. The detailed definition of CTPN can be found in A-C, Huang[13].

The model including Routing Module, Capability Group Module and Equipment Behavior Module are described in the following sections. Figure 1 illustrates the conceptual architecture of the proposed model.

Routing Module

The main purpose of Routing Module is to model the logical process flow of the wafer probing systems.

Each process step of the test flow has a processing capability (or operation type) associated with it. In the sort stage, each equipment may have more than one processing capability. Thus, when we classify these equipments into groups by processing capability, different processing capability groups may have equipments in common.

Figure 2 illustrates the detailed CTPN of Routing model. In this module, tokens, i.e. lots, having initial color attributes enter place Enter_P, and after firing mapping transition Enter_T enter place NextStage_p. The attribute will be changed for presenting lot's status. The token will enable one of transitions capabilityi_T (i = 1,...,n) by token's color attribute. After firing this transition, lot will enter the capability group module which defines correct equipments for this product. After finishing current step, token will go out from this capability group to one of places Lowyield_P and Out_P by sorting result. If lot enters place Lowyield_P, it means this lot has some problem and will enter place Out_P after engineer solves it. We use a stochastic transition with gamma distribution to present the solving time. Token in place Out_P will fire mapping transition Step_T and Step_T changes it's attribute to present that it enter the next step. This cycle will not be terminated until lot finished its all steps.

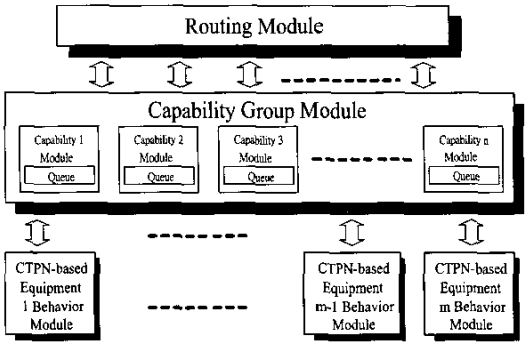

Figure 1 Conceptual architecture of the wafer probe center

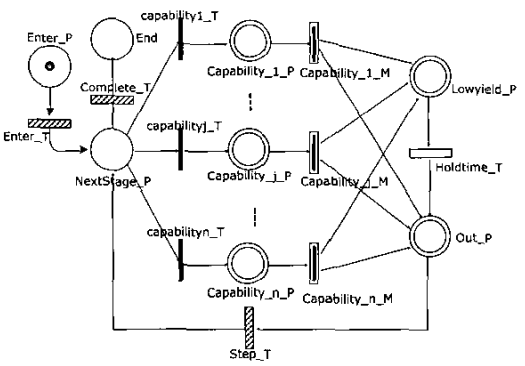

Figure 2 Routing module

Capability Group Module

The capability module contains one or more identical machines (or processing units). This module is responsible for dispatching the lot to an appropriate equipment behavior module.

The processing steps after a token enters this module are described as follows. First, the lot enters the place "Capability_i_P" and wants to select an idle equipment. The main problem here is: how to decide the equipment that the lot should enter if idle equipments do exist? We solve this problem by applying various selection rules, called Equipment Selecting

Rules (ESR). We can use one of these rules in any favorable manner, or we can use the GA search described in Chapter 4 to find the appropriate rule. If there is no idle equipment, the lot will enter queue ("Queue_P") and wait for selection by the equipment which just finishes operation. Figure 3 illustrates this module.

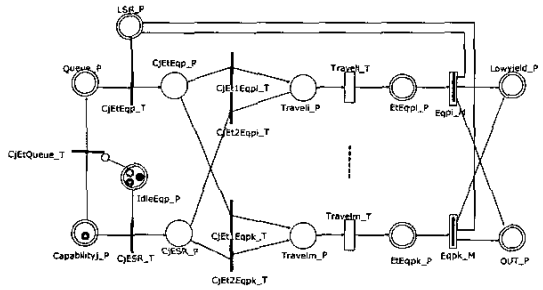

Figure 3 Capability group module

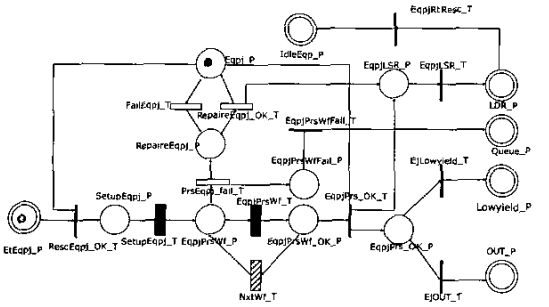

Figure 4 Equipment behavior module

Equipment behavior module

The purpose of the Elementary Module is to model the detailed equipment behavior in a wafer probe center, such as processing, setup, unscheduled equipment breakdown, and lot select operation. If the lot finishes testing, it will go into *place OUT_P* or *Lowyield_P* according to the testing result and the equipment will trigger the lot selection rule to select a lot in queue for the next operation. However, if the equipment encounters breakdown during processing, the unfinished lot will be sent into place *Queue_P* without changing its attribute which indicates the first unfinished wafer number, and when this lot resumes its operation, it will be processed from the first unfinished wafer number recorded last time. Figure 4 illustrates this module.

4. Wafer Probe Center Predictor and Scheduler

In order to provide more information to the customer, and help manager to make ordering decisions, we need to estimate some performance measures precisely and efficiently. In this section, we use the detailed wafer probing model described in Section 3 to predict the delivery date of the order and schedule the lot's ordering.

4.1 Wafer Probe Center Prediction

Since CTPN is one of semi-simulation tool, we can construct our system model and simulate it. Thus, there are some parameters needed initially to simulate the system. We describe these parameters in the following items.

Current system state: This information includes the current status of the lots, accessories, and the equipments.

Future Demands: These are the potential orders, which will be released in the future.

Yield Parameters: These are Yield information of each product.

Dispatching rules: These dispatching rules used now and will be adopted in the future.

Other parameters: These parameters include the probability of equipment failure, repair time of equipment, and some other minor parameters that will affect the precision.

The predictor can also help to answer some questions by changing some parameters and simulating result.

4.2 Wafer Probe Center Scheduling

It is important to develop a production scheduling system that helps to minimize cycle time, WIP, and to maximize equipment utilization as well as to meet the target constraints given by the manager. In our scheduler, we combine two viewpoints that one stands from lot and the other stands is from equipment. That means that lot can choose a fittest equipment based on current situation of these equipment by some rules, called equipment selection rules (ESR), and equipment can pick a fittest lot with current status of lots by some rules, called lot selection rules (LSR). Then we quantify these priority based rules and use Genetic Algorithm to find optimal solution.

Equipment Selection rule

The equipment selection rule is used to select an idle equipment. After a lot moved out from its last step, it enters the capability group corresponding to the current step. This selection must be considered from the lot's aspect. It must consider the status of this lot and all the idle equipments. In a probe center, testing a product will need to select an equipment with a particular test program, a probe card, and a load board. By this viewpoint, we need to think about the testing time resulting from the combination of test program, equipment, and probe card. Besides that, equipment information like utilization can also be taken as decision factor. Choosing the suitable equipment among all the possible ones to perform a processing step is definitely a key to the performance.

Lot Selection Rule

Lot selection rules hold the key position in controlling the performance of the wafer probing. It refers to how the priority value is assigned to jobs waiting in the queue. This selection should take the consideration from the equipment's view. It must consider the status of the queueing lots. Which queueing lots can be tested on this equipment without changeover of the accessories? Which combination of accessories has the shortest testing time or setup time? Which queueing lot is urgent? Which queueing lot has been queued for a long time? Which production type has a large quantity deficiency from the scheduled target? Thus, how to choose a queueing lot among all possible ones to process on equipment is certainly also a key issue to the performance.

Genetic Algorithm

There are two sub-problems in the scheduling problem. First, what searching method do we apply to optimize the scheduling results? Next, how do we evaluate the performance of each scheduling policy? For the first sub-problem, we apply a Genetic Algorithm. For the second sub-problem, we use simulation based on our CTPN model to complete the performance evaluation. In other words, we evaluate the performance indices in our fitness function through analysis of the proposed CTPN model.

Mixed Dispatch Rule

The main idea of this scheduler in this paper is to use a GA algorithm to find the appropriate equipment selection rule and lot selection rule. In addition, we try to mix some dispatching rules together and use GA search to find the appropriate mixed mode of each selection. This method, first normalize the value of priority-base rules used in ESR and LSR into the same range. Then, it uses different combinations of these ESR rules to generate different priority sequence of equipment for lots, and different combinations of these LSR rules to generate different execution sequence of lots for equipments. The detailed normalization method can be found in Huang[13].

Chromosome Representation

There are two types of genes in our chromosome: g_e, and g_l, which denotes equipment selection rule and lot selection rule, respectively. The contents of equipment selection rule $g_e = (a,b,c)$ is a three-tuple where a, b, and c stands for the selecting weight for UTIL_L, MTT, MST rules.

The contents of lot selection rule $g_l = (a,b,c,d,e,f)$ is a six-tuple where each variable respectively stands for six lot selection rules listed below.

For equipment selection rules in g_e, they are described simply as follows:

UTIL_L : Lowest Utilization

MTT : Minimum Testing Time

MST : Minimum Setup Time,

For lot selection rules in g_l, they are described simply as follows:

MTT : Minimum Testing Time

MST : Minimum Setup Time

MQT : Maximum Queueing Time

EDD : Earliest Due Date first

LST : Least Slack Target

Penalty

The rules we selected for gene codes are listed in Table 1.

Table 1 Gene codes

Equipment Selection Rule		*Lot Selection Rules*	
Name	Code	Name	Code
MTT	0 ~ 9	MTT	0~9
MST	0 ~ 9	MST	0~9
UTIL_L	0 ~ 9	MQT	0~9
		EDD	0~9
		LST	0~9
		Penalty	0~9

The gene codes are directly related to our CTPN Model. First, the ESR is related to the transition *CjESR_T*. When a token enters the place *Capabilityj_P*, it enables and fires transition *CjESR_T*. The result of firing is determined according to the selected value of g_e. The result of equipment selection is represented in color of the token in place *CjESR_P*, and this token will follow the specific color to enable the transition *CjEt2Eqpk_T*. The transition *EqpjLSR_T* is related to the LSR. When there is a token in place *EqpjLSR_P* and the transition *EqpjLSR_T* is enabled, we select the token (lot in place *Queue_P*) with the highest priority value to be tested in the j-th Equipment.

Fitness Function

In common situations, we have multiple objective measures. We use weighted-sum approach to combine them into a scalar fitness solution. The definition is as follows: (c is a chromosome (i.e., solution))

$$f(c) = w_1 \cdot f_1(c) + \cdots + w_5 \cdot f_5(c)$$

where f_i are the scores for mean cycle time, meet-target-rate, throughput rate, the loss due to orders which failed to meet due date, and pure utilization.

These scores are obtained from the CTPN model of the wafer probe center. We can encode any possible solutions into the chromosomes. Then we obtain the quality of the individual chromosome through simulation on the PN platform to guide the genetic evolution to produce a good schedule.

5. Experiment

To verify the effectiveness of the proposed work, we build a virtual wafer probe center model based on the data from a semiconductor fab in Hsinchu, Taiwan.

5.1 Experiment Specification

In this model, there are totally 122 equipments with 20 equipment types and 64 capabilities. The number of equipment of each capability may range from 1 to 10. There are 105 products with 105 routes in this model. There are 123 load boards and 330 probe cards with dots from 1 to 12. The probe card can be used for 1~5 kinds of product. The testing time is from 5 minutes to 412 minutes for each wafer and the setup time is from 6 minutes to 206 minutes. The testing time of each part is dependent on equipment type, stage, and number of dots in probe card, and the setup time of each part is dependent on equipment. Since the number of equipments in wafer probe center and routes of products is large, the complexity in simulation and scheduling is dramatically increased.

We assume that lots queueing in ovens will be picked randomly until batch capacity is full when the oven is open. The batch size of each oven is up to 12 lots. Finally, the parameters such as MTTF (Mean Time to Failure) and MTTR (Mean Time to Repair) are also included in this model, and we assume the MTTF is random from 100 hours to 200 hours and MTTR is from 1 hour to 5 hours.

5.2 Experiment Results

In these experiments, the wafer probe center is not empty in the beginning of evaluation. This means that there are lots which are already in processing in the wafer probe center. We generate the information of these lots by taking a snapshot of the simulation model already in a steady state. The period of evaluation is 15 days if we do not specify it later. There are three cases in our experiment. The first case is to compare our method with LSR approach. The second case is to compare our method with ESR approach. The final case it to compare our method with Huang's method [13]. The experiment will try to compare mean cycle times, penalty, WIP, meet target rates, mean throughput rates, and pure utilization. The compared result is given in following subsection.

ESR Approach

In the ESR approach, the lot picks an equipment using ESR when the lot enters a process stage. After picking the suitable equipment, the lot will be queued in its queue, and the equipment will pick a lot in queue of this equipment randomly when it becomes available. The ESR used in this approach is a mixture of of SQL (Shortest Queue Length), MTT, MST, and UTIL_L rule. Figure 5 shows the comparison of these two approaches.

LSR Approach

The LSR approach is similar to our approach. The difference is that the lot will randomly pick an idle equipment without using any rule when there are idle equipment in the capability group. This approach focuses on selecting lots when the equipment becomes idle. In this approach, LSR uses mixed LSR rule described in section 4. Figure 6 shows the comparison of these two approaches.

Huang's Approach

The Huang's approach [13] is similar to ESR approach. In this approach, equipment has a queue and lot will pick an equipment using MSR (namely Machine Selection Rule in Huang's method) when the lot enters capability. After equipment has tested one lot, it will pick a lot in queue of this equipment using DR (Dispatching rule). The MSR used in this approach is a mixture of SQL (Shortest Queue Length), MTT, MST, and UTIL_L rules. The DR uses mixed LSR rules, which is the same as our lot selection rule. Figure 7 shows the comparison of these two approaches.

5.3 Analysis of Experiment Results

From Figure 4, we found that the proposed GA scheduler performs much better than ESR approach. It has lower penalty, WIP, pure utilization, higher target rate and throughput rate. Despite that ESR approach performs better than our approach in mean cycle time, our approach reduces penalty, raises target rate and throughput rate, and also reduces WIP level by more than half. In Table 6, experiment shows that the proposed GA scheduler performs much better than LSR approach. It has lower mean cycle time, penalty, pure utilization, higher target rate, throughput rate, and a lower WIP level. In Figure 7, experiment also shows that the proposed GA scheduler performs better than Huang's approach. It has lower mean cycle time and penalty, higher throughput rate, and a lower WIP level. Despite that Huang's approach performs better than our approach in target rate and pure utilization, our approach reduces more mean cycle time and penalty, increases more throughput rate, and decreases WIP level by more than one seventh. As a result, the proposed GA scheduler has a significant improvement on key performance of wafer probe center.

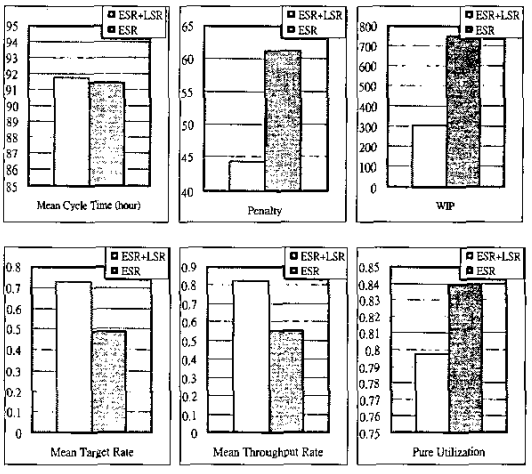

Figure 5 Comparison of our approach and ESR approach

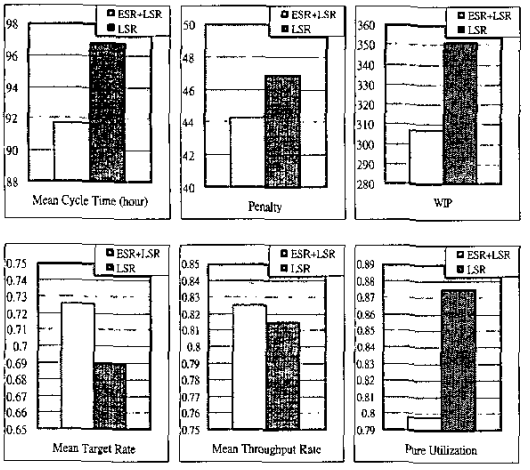

Figure 6 Comparison of our approach and LSR approach

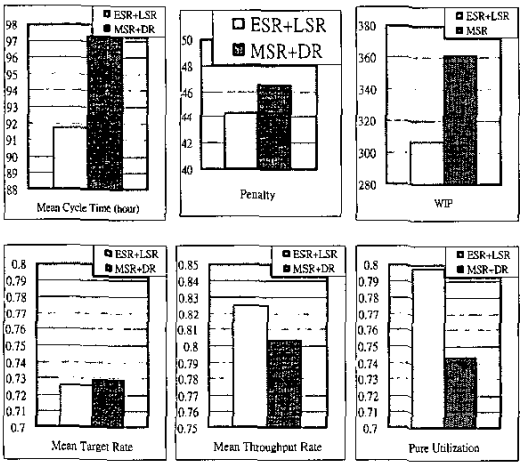

Figure 7 Comparison of our approach and LSR approach

6. Conclusion

In this paper, we propose a wafer probe center model based on colored timed Petri-net. We can predict due-date of any specific product under some scheduling policies by simulation on the proposed CTPN model. In scheduling phase, we consider ESR when lots enter a capability module with existing idle equipment, and LSR when an equipment finishes testing and has more than one lot in queue. This approach can be extended more widely. First, let the lot select the equipment and generate a selection sequence with considerations of the factors of each individual equipment. Second, let the equipment generate sequence of lots standing from the viewpoint of the equipment. After generating these two categories of sequences, we can then compare these two sequences and find the best matching of lot and equipment using defined evaluation function. Since much more information is used at the decision point, it will increase solution space dramatically. We can also apply GA to these approaches and find the optimal solution. Because of the large solution space, GA can find more effective rules. The extension of our method will be left as our future work.

Reference

[1] M. H. Lin and L. C. Fu, "Modeling, Analysis, Simulation, and Control of Semiconductor Manufacturing Systems: A Generalized Stochastic Colored-Timed Petri-Net Approach," *IEEE International Conference on Systems Man, and Cybernetics*, 1999.

[2] S. Y. Lin and H. P. Huang, "Modeling and Emulation of a Furnace in IC Fab Based on Colored-Timed Petri Net," *IEEE Transactions on Semiconductor Manufacturing*, vol. 11, no. 3, pp. 410-420, 1998.

[3] H. H. Xiong and M. C. Zhou, "Scheduling of Semiconductor Test Facility via Petri Nets and Hybrid Heuristic Search," *IEEE Transactions on Semiconductor Manufacturing*, vol. 11, no. 3, pp. 384-393, 1998.

[4] M. Allam and H. Alla, "Modeling and Simulation of an Electronic Component Manufacturing System Using Hybrid Petri Nets," *IEEE Transactions on Semiconductor Manufacturing*, vol. 11, no. 3, pp. 374-383, 1998.

[5] C. Y. Lee, R. Uzsoy, and L. A. Martin-Vega, "Efficient Algorithms for Scheduling Semiconductor Burin-in Operations," *Operation Research*, vol. 40, no. 4, pp. 764-795, 1992.

[6] S.-C. Huang and J. T. Lin, "An Interactive Scheduler for A Wafer Probe Center in Semiconductor Manufacturing," *International Journal of Production Research*, vol. 35, no. 4, pp. 1171-1191, 1997.

[7] Y. D. Kim, J. U. Kim, S. K. Lim, and H. B. Jun, "Due-Date Based Scheduling and Control Policies in a Multi-product Semiconductor Wafer Fabrication Facility," *IEEE Transactions on Semiconductor Manufacturing*, vol. 11, no. 1, pp. 155-164, 1998.

[8] T. Murata, H. Ishibuchi, and H. Tanaka, "Multi-Objective Genetic Algorithm and Its Applications to Flowshop Scheduling," *Computers and Industrial Engineering*, vol. 30, no. 4, pp. 957-968, 1996.

[9] R. Sikora, "A Genetic Algorithm for Integrating Lot-Sizing and Sequencing in Scheduling a Capacitated Flow Line," *Computers and Industrial Engineering*, vol. 30, no. 4, pp. 969-981, 1996.

[10] R. Cheng, M. Gen, and Y. Tsujimura, "A Tutorial Survey of Job-Shop Scheduling Problems Using Genetic Algorithm, Part I: Representation," *Computers and Industrial Engineering*, vol. 30, no. 4, pp. 983-997, 1996.

[11] A. C. Huang, L. C. Fu, M. H. Lin, and S. Y. Lin, "Modeling, Scheduling, and prediction for wafer fabrication: queueing colored Petri-net and GA based approach," Robotics and Automation, 2002. Proceedings. ICRA '02. IEEE International Conference on , Volume: 3 , 2002 Page(s): 3187 -3192 vol.3