

This article was downloaded by: [National Taiwan University]

On: 12 January 2009

Access details: Access Details: [subscription number 732449060]

Publisher Taylor & Francis

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



International Journal of Production Research

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title-content=t713696255>

Reverse auction-based job assignment among foundry fabs

Shi-Chung Chang ^a; Ming-Ming Hsieh ^a; Chia-Wei Chen ^a

^a Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan

Online Publication Date: 01 February 2007

To cite this Article Chang, Shi-Chung, Hsieh, Ming-Ming and Chen, Chia-Wei(2007)'Reverse auction-based job assignment among foundry fabs',International Journal of Production Research,45:3,653 — 673

To link to this Article: DOI: 10.1080/00207540600792176

URL: <http://dx.doi.org/10.1080/00207540600792176>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Reverse auction-based job assignment among foundry fabs

SHI-CHUNG CHANG*, MING-MING HSIEH and CHIA-WEI CHEN

Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan, R.O.C 10617

(Revision received March 2006)

Motivated by foundry service provisioning in the semiconductor industry, this paper adopts a reverse auction-based mechanism to model job order assignment from a job owner (auctioneers) to a few qualified and competing foundry fabs (bidders). A job order owner announces job requirements and payments for fabs to bid on while qualified fabs bid on a job by offering the discount to payment and processing schedule of the job. This model aims at capturing the non-cooperative gaming among the job owner and the fabs because of the private information such as objectives, valuation of jobs, available capacity and constraints. Two integer-programming formulations are formulated for the reverse auction of job assignment; one is deterministic and models a common practice of mean value-based bidder decision-making while the other is stochastic and captures bidders' consideration of uncertainties and the associated risk in decision-making. A Lagrange relaxation-based, near-optimal scheduling method is developed to model a bidder's selection of job-to-bid and schedule for the deterministic formulation. The bidder's job selection and scheduling model for the stochastic formulation combines simulation and simple heuristics. A bidder decides the discount offer for each job-to-bid by a simple fixed-increment scheme. The auctioneer simply assigns the job to the bid with the highest discount offer in each round of bidding. Analyses show that the reverse auction model leads to an equilibrium solution, where no single bidder would unilaterally deviate from the auction result. Numerical study by using the reverse auction model demonstrates that consideration of uncertainties in bidders' decisions has a larger impact on performance of both bidders and the auctioneer than optimality of the mean value-based bidder scheduling algorithm.

Keywords: Foundry manufacturing; Reverse auction; Modelling; Assignment; Scheduling; Lagrange relaxation

1. Introduction

Foundry service is a significant part of contract manufacturing in the semiconductor supply chain (Su *et al.* 2005). The revenue of foundry business has seen a fast growth in the past decade and is now approximately 20% of the semiconductor wafer fabrication. On the one hand, fabless design houses of integrated circuits generate various job orders and compete for the capacity of qualified foundry service providers. On the other hand, foundry fabs/companies compete to get the job orders by providing low-cost and timely manufacturing services. To match customers' needs

*Corresponding author. Email: scchang@cc.ee.ntu.edu.tw

with chip fabrication processes, foundry brokers have recently emerged that forge alliances with foundry fabs and serves as a 'value-added broker' (Clarke 2004). Fabs within a foundry company may also compete internally for job orders from the central planning.

In a competitive foundry service environment, non-cooperative gaming phenomena naturally exist among all the decision-makers (DMs), i.e. job owners and fabs. Individual DMs have different domains of management and control, various objectives, and private information such as ones' own valuation of jobs, actual capacity and operational constraints. How the job orders are assigned or scheduled to be manufactured by which qualified fabs at what cost must be done via a distributed decision mechanism with significant gaming considerations. It is therefore very challenging to model and design the job assignment/scheduling mechanism for semiconductor supply chain management.

Models for job assignment and scheduling have a long history in supply-chain management (Fan *et al.* 2003, Luh *et al.* 2003). However, most of the earlier articles in the literature have been on centrally managed supply chains. There has been a growing literature that examines decentralized chains, usually with information asymmetry. A great deal of attention has been paid to the auction-based job scheduling or assignment (Bertsekas 1990, Dewan and Joshi 2002, Luh *et al.* 2003, Sarne *et al.* 2004, Srivinas *et al.* 2004). There are a few salient features in an auction market. First, the value of the merchandise is not obvious and not the same to DMs. Second, the buyers are independent and seek their own profits. Third, they also have their own private information such as budget constraints and the valuations of merchandises. Most of recent research considered an auction market with one auctioneer and several bidders. In each round of bidding, the bidders send job 'bids' to the auctioneer and the auctioneer then temporarily assigns each job to the bidder who offers the highest bid on the job. The highest bid on each job serves as the starting bid for the job in the next round of bidding. The auction repeats round by round until no new bids are offered and the final assignment is determined. Except the job information of processing and delivery requirements and bidding information of price, starting date and delivery date, there is no need for DMs to know the objective functions and private information amongst each other.

In the literature, there have been extensive surveys of auction mechanisms (Milgrom and Weber 1982, de Vries and Vohray 2003). Bertsekas (1990) designed an auction algorithm to solve various assignment and network optimization problems. The emphasis was on exploiting the parallelism of an auction-based scheme instead of the gaming aspect. Kutanoglu and Wu (1999) studied a distributed scheduling framework by exploiting Lagrangian relaxation-based decomposition and coordination. The basic idea is to localize and distribute the functionality of operational scheduling decisions. Each DM has a local problem to maximize his/her expected total reward subject to local constraints. The result can then be communicated to the coordination mechanism as a 'bid'. The coordination mechanism or an 'auctioneer' is a bid processor that makes resource allocation based on an iterative auction process using the bid information.

In the context of contract manufacturing, Karabuk and Wu (2002) examined issues of a decentralized coordination scheme between capacity reservation by marketing managers and capacity allocation to competing marketing managers by manufacturing managers. The authors compared the gap between centralized and

decentralized allocation models, evaluated the price of decentralization and proposed an information exchange principle. Karabuk and Wu (2005) further modelled a semiconductor capacity-allocation problem in a game-theoretic setting and designed an incentive scheme through bonus payments and participation charges that elicit private demand information. They showed that their mechanism achieves budget-balance and voluntary participation requirements simultaneously. Their results provide important insights into the treatment of misaligned incentives.

In deregulated electricity markets, where electricity supplies cannot be specifically assigned to load demands, the market clearing price (MCP) auction mechanism has been a common practice (Hao 2000, Hogan 2005). Under the MCP auction mechanism, market participants submit supply offers and demand bids for energy and ancillary services to an independent system operator (ISO). The ISO then runs auctions for energy and ancillary services to minimize the total payment and determine the MCPs for each product. To determine MCP price and supply offers in a sealed-bid pay-at-MCP problem, which is characterized by non-separability in the objective function and discontinuity in offer curves, Luh *et al.* (2006) combined the augmented Lagrangian relaxation approach with an advanced surrogate optimization framework into an effective solution method.

This paper addresses the modelling of job order assignment from a job owner to a few qualified and highly competitive foundry fabs. The foundry service provision conditions of one-buyer (job owner) and multiple-sellers (foundry fabs) structure, competitiveness among foundry fabs for orders from design houses, and clear specifications of foundry service required by each job make a reverse auction-based mechanism (de Vries and Vohray 2003, Fan *et al.* 2003) favourable for modelling the job competition and assignment decisions (Ghawi and Schneider 2004). A job order owner announces job requirements and payments for fabs to bid on while qualified fabs bid on a job by offering the discount to payment and the processing schedule of the job. This model aims at capturing the non-cooperative gaming among the job owner and the fabs because of the private information such as objectives, valuation of jobs, available capacity and constraints. Two integer-programming formulations are formulated to represent two types of bidders' logistic decision capability; one is deterministic and models a common practice of mean value-based decision-making, while the other is stochastic and captures the consideration of uncertainties and the associated risk in decision-making.

In addition to mathematical problem formulations, this paper develops the corresponding decision algorithms for individual DMs as part of the model. Given the bid information from the auctioneer, a bidder's selection of job-to-bid and scheduling for the deterministic formulation are based on a near-optimal approach of Lagrange relaxation (Hoitomt *et al.* 1993, Tang *et al.* 2002). The bidder's job selection and scheduling model for the stochastic formulation combines simulation and simple heuristics. A bidder then decides the discount offer for each job-to-bid by a simple fixed-increment scheme. The auctioneer simply assigns a job to the bid with the highest discount offer in each round of bidding.

Properties of the reverse auction model such as the existence of an equilibrium and convergence in finite rounds are analysed. The reverse auction model is then used to assess the effects of problem formulation and scheduling algorithm. The study compares job assignments obtained by using three different combinations of formulation and scheduling algorithm: LR-based algorithm for the deterministic

formulation (D-LR), simulation and heuristic-based scheduling algorithm for the deterministic formulation (D-SH), and simulation and heuristic-based scheduling algorithm for the stochastic formulation (S-SH). Numerical experiments with the three combinations demonstrate that S-SH leads to lower discount offers and higher bidder profits than the other two. Such results imply that consideration of uncertainties in bidder's decision has a larger impact on performance of both bidders and the auctioneer than optimality of the mean value-based scheduling algorithm that a bidder adopts.

The remainder of the paper is organized as follows. Section 2 first gives an integer programming problem formulation of the reverse auction problem. There are two formulations for bidder's problem: deterministic and stochastic. Section 3 then develops the complete and iterative solution algorithms, which include (1) a Lagrange relaxation-based and a simulation and heuristic-based scheduling algorithms for bidder job selection and scheduling, (2) a simple fixed increment scheme for bidder discount offering and (3) the best-bid job assignment by the auctioneer. Section 4 then presents solution analysis and numerical studies of the reverse auction model. Finally, section 5 concludes the paper.

2. Problem formulation

Consider a foundry service environment of semiconductor manufacturing, where there is a job provider who offers a set of jobs and calls for bids from a few fabs to process the jobs. This paper assumes a simplified problem, where each job has only one operation and requires a certain processing time of a unit capacity from a foundry service provider. The job processing is non-preemptive; that is, once a job is started, no interruptions are allowed until completion. Therefore, a job can only be assigned to one fab at a time and there are no inter-fab transfers or backups. Each job has a time window for processing between its target release date and due date. If a job is planned to start before the target release date or to be delivered after the due date, there will be a penalty incurred on the foundry service provider. Assume that jobs under consideration are already available for processing.

The job provider out-sources the job processing to qualified fabs and sets for the processing of each job a maximum payment. The fabs are competitive to each other. Each fab has its own private information such as the actual capacity, the job processing cost, the valuation of getting a job and its objective function. Without knowing such private information of fabs, the job provider cannot perform centralized job scheduling; the scheduling has to be decentralized.

As the foundry service provision conditions of one job owner and multiple fabs, competition among fabs, and clear specifications of job requirements make reverse auctions favourable for modelling job competition and assignment (Ghawi and Schneider 2004), in this paper, we consider a reverse auction-based mechanism (de Vries and Vohray 2003, Fan *et al.* 2003). The objective of the job provider (auctioneer) is to minimize its payment cost of completing all the jobs in hand plus the penalty cost of unassigned jobs, if any. By distributing job information to all the qualified fabs (bidders), the job provider solicits individual fabs to bid on the jobs, where a bid on a job specifies the beginning time, the completion time and the discount from the maximum payment for processing the job. In each round of

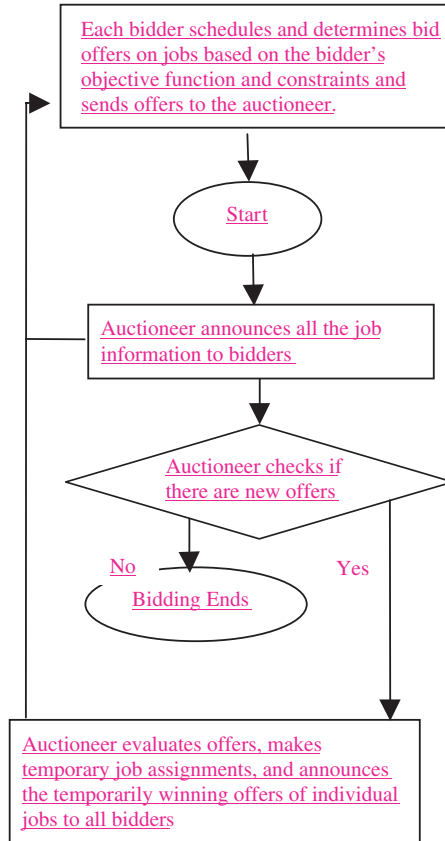


Figure 1. Flowchart of a reverse auction.

bidding, the job provider first selects the fab that offers the highest discount of each job as the temporary assignment of the job. The initial discounts are zero for all jobs.

Given the temporary discounts on assignments of jobs, a fab then evaluates and schedules to determine whether to offer new bids on various jobs. The new discount offer for a job must be higher than its current value. The objective of a fab is to maximize the payment it may receive from processing the jobs minus the earliness/tardiness penalty for each job not processed within the desired time window of the job. Note that the earliness/tardiness penalty of a job is fully compensated by the job winning fab. After the job provider collects the bids from all the fabs, one round of bidding ends. Such a procedure repeats round by round until no new bids from any fabs. The flowchart of a reverse auction is in figure 1, which provides a solution framework to the non-cooperative game among the job provider and fabs.

2.1 Auctioneer's decision problem

The objective of the auctioneer is to have the jobs completed on time at as low a cost as possible. Since the auctioneer does not have the private information of individual

Table 1. Notations.

i : Job index, $i = 1, \dots, N$	t : Time slot index, $t = 1, \dots, T$, where T is the time horizon under consideration	k : service provider (fab) index, $k = 1, \dots, M$
v_i : Maximum payment to processing job i	f_i : Penalty on job i if i is not assigned	λ'_{ik} : Offered discount of job i by fab k
λ_i : Discount for job i	ES_i : Earliest starting date	r_i : Earliness penalty of job i
d_i : Tardiness penalty of job i	R_i : Release date of job i	D_i : Due date of job i
γ_i : Minimum tardiness penalty	ρ_{ik} : Operation cost of fab k for processing job i	p_{ik} : Processing time of job i by fab k

bidders, the auctioneer assigns a job to a fab based on bidders' bids. Let us first define some notations in table 1.

Define the auctioneer's assignment decision variable as:

$$X_{ik} = \begin{cases} 1, & \text{if job } i \text{ is assigned to bidder } k; \\ 0, & \text{otherwise;} \end{cases} \quad \forall i, k; \quad (1)$$

and a bid indicator variable based on bidders' bids as:

$$\bar{B}_{ik} = \begin{cases} 1, & \text{if bidder } k \text{ bids on job } i; \\ 0, & \text{otherwise;} \end{cases} \quad \forall i, k; \quad (2)$$

where both types of variables are integral variables. A job can be assigned to at most one bidder. i.e.:

Single assignment constraint

$$\sum_k X_{ik} \leq 1, \forall i, \quad (3)$$

and a job can be assigned to a bidder only when the bidder bids on the job, i.e.:

Bid-based assignment constraint:

$$X_{ik} \leq \bar{B}_{ik}, \forall i, k. \quad (4)$$

The payment discount of a job is the discount offered by the bidder assigned to the job:

Job payment discount equation

$$\lambda_i = \sum_k X_{ik} \lambda'_{ik} \quad \forall i. \quad (5)$$

The objective function consists of the payment cost of completing all the jobs in hand plus the penalty cost of unassigned jobs. Combined with the constraints above, the auctioneer's decision problem (ADP) is:

$$\text{(ADP) Minimize}_{\{X_{ik}\}} \sum_i \sum_k X_{ik} (v_i - \lambda_i) + \sum_i \left(1 - \sum_k X_{ik} \right) f_i \quad (6)$$

Subject to constraints (1) to (5).

2.2 Bidder's Decision Problem

When bidders receive the job and bid information, they must decide which jobs to bid on at what discount offer, when to begin and complete processing. Bidders compete for jobs by offering discounts to the payment of job processing; the higher the discount, the lower the job processing cost to the auctioneer but the lower profit margin to the bidder. Once a bidder is assigned a job, the bidder is responsible for delivery within the desired time window and has to take the penalty for earliness or tardiness in delivery by contract. A bidder's objective is to maximize the bidder's profit, which equals to the payments from processing the winning jobs minus earliness/tardiness penalties if any.

2.2.1 Deterministic formulation. In this formulation, a bidder uses the mean processing time to schedule jobs and determine bids. Define the decision variables of a bidder k as

$$B_{ikt} \equiv \begin{cases} 1, & \text{if bidder } k \text{ will start processing job } i \text{ at time } t; \\ 0, & \text{otherwise;} \end{cases} \quad \forall k, t. \quad (7)$$

Let $\{C_{kt}, \forall t\}$ be the capacity over time, which is private information of bidder k . Assume that a fab bids within its capacity, which means the number of jobs bid does not exceed the capacity at any time. A bidder bids on a job at most once.

Capacity constraint:

$$\sum_i \sum_{t=\bar{p}_{ik}+1}^t B_{ikt} \leq C_{kt}, \quad \forall k, t, \quad (8)$$

where $\bar{p}_{ik} = E[p_{ik}]$, $\forall i, k$.

At most one beginning time per job constraint:

$$\sum_t B_{ikt} \leq 1, \quad \forall i. \quad (9)$$

In order to be accepted by the auctioneer, a new discount offer on a job must be higher than the currently announced discount for the job.

Increasing offer constraint:

$$\lambda'_{ik} > \lambda_i, \quad \forall i. \quad (10)$$

As a job i is not available before its earliest starting date ES_i , the bid on a job cannot go before the available time of the job. So,

Earliest starting time constraint:

$$\sum_i B_{ikt} \cdot ES_i \leq t, \quad \forall t. \quad (11)$$

In our problem, we assume that the earliest starting date is one for all the jobs.

The objective function of a bidder consists of two terms. The first is from processing a job and equals the maximum payment minus the offered discount and

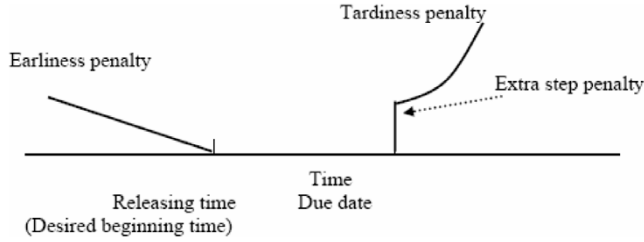


Figure 2. Penalty function of a job.

the cost of job processing. The second term corresponds to the earliness and tardiness penalty as shown in figure 2 (Yu *et al.* 2005). Note that there is a jump in the tardiness penalty once the delivery is late. A deterministic formulation of bidder's decision problem (*BDP/D*) is then:

$$\begin{aligned}
 (BDP/D_k) \quad & k = 1, \dots, K \\
 \max_{B_{ikt}, \lambda'_{ik}} \{ & U_k(B_{ikt}, \lambda'_{ik}) \equiv \sum_t \sum_i [B_{ikt}(v_i - \lambda'_{ik} - \rho_{ik})] \\
 & - \sum_t \sum_i B_{ikt} [r_i \max(0, R_i - t) + d_i \max(0, (t + \bar{p}_{ik} - D_i))^2 \\
 & + \gamma_i \text{Step}(t + \bar{p}_{ik} - D_i)] \} \quad (12)
 \end{aligned}$$

Subject to (2) and (7)–(11).

2.2.2 Stochastic formulation. This formulation assumes prudent bidders, who take into account the uncertainties of processing time and the risk incurred by the uncertainties. Decision variables of a bidder, say *k*, are still $\{B_{ikt}, \forall_{i,t}\}$ as defined in equation (7) and bidder *k*'s capacity over time, $\{C_{kt}, \forall_t\}$, is private information to the bidder. Instead of using the expected processing time for scheduling, a bidder schedules with risk consideration of individual jobs under job processing time uncertainty. A bidder's decision problem is then:

$$\begin{aligned}
 (BDP/S_k) \quad & k = 1, \dots, K \\
 \max_{B_{ikt}, \lambda'_{ik}} \{ & U_k(B_{ikt}, \lambda'_{ik}) \equiv E \left\{ \sum_t \sum_i [B_{ikt}(v_i - \lambda'_{ik} - \rho_{ik})] \right. \\
 & - \sum_t \sum_i B_{ikt} [r_i \max(0, R_i - t) + d_i \max(0, (t + p_{ik} - D_i))^2 \\
 & \left. + \gamma_i \text{Step}(t + p_{ik} - D_i)] \right\} \} \quad (13)
 \end{aligned}$$

subject to (9)–(11) and

$$\sum_i \sum_{t=p_{ik}+1}^t B_{ikt} \leq C_{kt}, \quad \forall k, t, \quad (8a)$$

where random job processing times are considered instead of the mean processing times as in equation (8).

Note that in (BDP/D_k) or (BDP/S_k) , if the solution has $\sum_t B_{ikt} = 1$, then fab k bids on job i , i.e. $\bar{B}_{ik} = 1$ in equation (2) of ADP. In the problem formulation above, bidders' problems (BDP/D_k) and (BDP/S_k) are much more complicated than the auctioneer's problem (ADP) because of the complexity of fab operations. Bidders' job-to-bid selection and scheduling require intensive decision power and would significantly affect the auction results.

3. Solution methodology

In both deterministic and stochastic formulations of BDP, a bidder's bid makes two types of decision: job-scheduling and discount-offering. This section takes a two-stage approach to develop a solution method for each problem. In the approach, a bidder first decides which jobs to bid on by solving BDP for the scheduling decisions based on the currently announced discounts, $\{\lambda_i\}$, of individual jobs. A bidder would schedule as many profitable jobs as possible based on the trade off between the potential gain $\{v_i - \lambda_i - \rho_{ik}\}$ and the earliness/tardiness penalty cost. The bidder then decides the discount offer for each job to bid on.

Once individual bidders' bids are available, the auctioneer's bid selection and job assignment are quite straightforward by following the highest bid assignment and one assignment per job rules. The auctioneer's solution algorithm is independent of the solution algorithm for BDP. Specific solution methods for (BDP/D) , (BDP/S) and (ADP) are now described as follows.

3.1 Lagrangian relaxation (LR)-based scheduling algorithm for (BDP/D)

We adopt a Lagrangian relaxation (LR)-based job shop scheduling approach developed by Hoitomt *et al.* (1993) to solve the scheduling problem of a bidder k . By relaxing the capacity constraint, Hoitomt *et al.* (1993) decomposed the original problem into simple and independent list scheduling sub-problems among jobs. To coordinate among subproblem solutions, Hoitomt *et al.* (1993) adopted a subgradient method to adjust the relaxation multipliers. This approach generates both a dual solution and a primal solution, where the dual solution serves as a lower bound on cost function value. The relative difference in cost function values between the primal and the dual solutions then gives an indicator of how close the primal solution is to the optimum. Hoitomt *et al.* (1993) demonstrated that their LR-based scheduling algorithm efficiently achieves near-optimal solutions for small to medium job shop scheduling problems and Liao *et al.* (1996) and Hwang and Chang (2003) successfully designed LR-based scheduling algorithms for fab production scheduling.

Let us define a few more notations and omit the bidder index k in the remainder of this section for the conciseness of presentation.

Notations

- b_i : beginning time of job i under the condition that $\sum_t B_{it} = 1$, i.e. the bidder selects job i to bid on, and therefore $b_i \equiv \sum_t t \cdot B_{it}$;
- c_i : completion (delivery) time of a selected job i , where $c_i = b_i + \bar{p}_{ik}$;

$ER_i \equiv \max(0, D_i - C_i)$, the earliness of delivering job i ; and
 $TN_i \equiv \max(0, c_i - D_i)$ the tardiness of delivering job i .

Observe that the coupling among jobs in (BDP/D) is through the capacity constraint (inequality (8)). We apply Lagrange relaxation and associate a Lagrange multiplier π_t with the capacity constraint of time t to form a Lagrangian as follows:

$$L \equiv \sum_t B_{it} \left\{ - \sum_i [v_i - \lambda'_i - \rho_i] + \left\{ \sum_i [r_i \cdot ER_i + d_i \cdot TN_i^2 + \gamma_i \text{Step}(TN_i)] \right\} \right. \\ \left. + \sum_t \pi_t \left(\sum_i \sum_{t-\bar{p}_i+1}^t B_{it} - C_t \right) \right\}.$$

A bidder's BDP-S problem can then be decomposed into individual job scheduling problems and form a dual problem.

$$(JSP_i) \quad i = 1, \dots, I.$$

$$L_i^*(\{\pi_t\}) \equiv \min_{\{B_{it}\}} L_i(\{\pi_t\}, \{B_{it}\})$$

subject to (7), (9), (10) and (11) with $\{\pi_t\}$ given and

$$L_i(\{\pi_t\}, \{B_{it}\}) \\ \equiv \sum_t \left\{ B_{it} \cdot [(v_i \cdot ER_i + d_i \cdot TN_i^2 + \gamma_i \text{Step}(TN_i)) - (v_i - \lambda'_i - \rho_i)] + \pi_t \sum_{t-\bar{p}_i+1}^t B_{it} \right\},$$

and the dual problem is

$$(D) \max_{\{\pi_t\} \geq 0} L \equiv - \sum_t C_t \pi_t + \sum_i L_i^*(\{\pi_t\}).$$

The solution method of the dual problem then follows that of (Zhao *et al.* 1999) as sketched in the following descriptions.

1. Initialize $\pi_t = 0, \forall t$, and set an iteration counter $n = 0$. Solve (JSP_{*i*}) $i = 1, \dots, I$, by searching through the admissible values of B_{it} , where there is at most one time t when $B_{it} = 1$. The optimal beginning time decision $\{B_{it}^*, \forall t\}$ of job i is the one yielding the lowest value of L_i . The computational complexity of this search procedure is linear with respect to the number of time slots T . Note that the earliest starting date constraint can be easily accommodated by requiring the non-zero B_{it} to lie between the earliest starting date of job i and $T - p_i + 1$.
2. Once $\{\{B_{it}^*, \forall t\}, i = 1, \dots, I\}$ is obtained, the Lagrange multipliers $\underline{\pi} \equiv \{\pi_t, \forall t\}$ are then updated according to a sub-gradient method for solving the dual problem (D),

$$\underline{\pi}^{n+1} = \underline{\pi}^n + \alpha^n \underline{g}(\underline{\pi}^n),$$

where $g(\underline{\pi})$ is the sub-gradient of L with respect to the Lagrange multipliers $\underline{\pi}$ and the step size α^n at the n th iteration is calculated by

$$\alpha^n = \gamma \frac{\bar{L} - L^n}{g(\pi^n)^T g(\pi^n)}, \quad 0 < \gamma < 2,$$

where \bar{L} is an estimation of the optimal solution, and L^n is the value of L at the n th iteration.

Although Zhao *et al.* (1999) proved that this method converges to the dual optimal, there is no guarantee of feasibility, i.e. some capacity constraints may be violated in the dual solution. The value of the dual problem L^* is a lower bound of the minimum, J^* , of the primal problem (BDP/D).

To construct a feasible schedule from the optimal dual solution, a greedy heuristic is adopted. The feasibility adjustment is done sequentially over time. If the capacity constraint is violated at time t , the heuristic calculates, for jobs with beginning time at time t , the individual weighted tardiness penalties if their beginning time is delayed to time $t + 1$. The heuristic then assigns jobs to the available capacity of time t in the descending order of their tardiness penalties. The remaining jobs are delayed by one time unit and are considered together with those jobs originally scheduled to start at time $t + 1$. Such a procedure repeats over the whole time horizon and leads to a feasible schedule. Once a feasible solution is obtained, the corresponding value of the objective function J is an upper bound of J^* .

Note that a job may not be profitable after the feasibility adjustment of the schedule. Once a feasible schedule is obtained, the bidder keeps only profitable jobs as the jobs-to-bid. Let the set of jobs-to-bid by the bidder be $S^B \equiv \{i \mid \sum_t B_{it} = 1, i = 1, \dots, I\}$. The discount offer of each job-to-bid must increase from one round to the next. We assume a least quantum increase, ε , uniformly across all the jobs and bidders. So a bidder will offer a new discount for a job-to-bid i by

$$\lambda'_i = \begin{cases} \lambda_i + \varepsilon, & \text{if } i \in S^B; \\ 0, & \text{otherwise.} \end{cases}$$

3.2 Simulation and heuristic-based scheduling algorithm for (BDP/S)

Each (BDP/S_k) problem is a challenging problem of stochastic dynamic optimization. To mimic industrial applications of simulation and heuristics, simple heuristics with job schedule evaluation by simulation are adopted in the bidder's scheduling algorithm. The algorithm is sketched as follows.

Algorithm sketch

Step 1: Consider all the jobs that are unassigned and have $v_i - \lambda_i - \rho_{ik} > 0$. Sequence these jobs in a descending order of their $v_i - \lambda_i - \rho_{ik}$ values. Let the ordered job set be S_0 and the iteration index $i = 0$.

Step 2: Set $S = S_i$. Schedule the jobs in S one-by-one according to the sequence. Each job is started at the earliest possible time, i.e. the completion time of the previous job. A simple and single run simulation then determines the profit value and

the completion time of the job being considered. If the profit is positive, the job stays in \mathbf{S} ; otherwise, the job is deleted from \mathbf{S} . Collect the final \mathbf{S} into the sequence set Ω .

Step 3: Exchange the sequence between a pair of adjacent jobs in \mathbf{S}_i and set it as \mathbf{S}_{i+1} . Set $i \leftarrow i + 1$. Go to step 2. The iteration repeats till all the two-exchanges of \mathbf{S}_0 are evaluated.

Step 4: Perform Monte-Carlo simulation to evaluate each of the job sequence \mathbf{S} in Ω . In each replication, the jobs in \mathbf{S} are scheduled one-by-one according to the sequence in the same way as in Step 2. Among all the job sequences in Ω select the sequence with the highest expected profit as the final sequence. Let it be \mathbf{S}^* .

Step 5: Sequentially determine the beginning and completion times, i.e. scheduling, of jobs in \mathbf{S}^* according to (13).

3.3 Auctioneer's simple assignment algorithm

The solution algorithm to (ADP) is simple. Let $\{\lambda'_{ik}, \forall i\}$ be the discount offers made by bidder k . The auctioneer makes a new but temporary job assignment by

$$X_{ik} = \begin{cases} 1, & \text{if } k \in \arg \max\{\lambda'_{ik'}, \forall k'\} \text{ and } i \in S_k^B; \\ 0, & \text{otherwise;} \end{cases} \quad \forall i, k.$$

If the assignment of the previous round belongs to the solution set of this round, the auction terminates. If not, further rounds are needed. When there is a tie in discount offers among bidders for an object, a random selection is used to break the tie. The discount offers $\{\lambda_i, \forall i\}$ are then updated based on the offers of the assigned bidders, i.e.

$$\lambda_i = \lambda_{ik}, \text{ where } X_{ik} = 1, \forall i.$$

4. Performance analysis and evaluation

Now investigate the properties of the reverse auction algorithm. First let us analyse, in the deterministic problem context, if the auction terminates in finite steps and the property of the solution. Then let us conduct numerical study to investigate how bidders' use of information and scheduling algorithms may affect job assignments via the reverse auction.

4.1 Convergence analysis

To examine the termination of the algorithm, let us consider the profit of bidder k by bidding on job i

$$pf_{ik} = v_i - \lambda'_{ik} - \rho_{ik} - [r_i \max(0, R_i - t) + d_i \max(0, (t + \bar{p}_{ik} - D_i))^2 + \gamma_i \text{Step}(t + \bar{p}_{ik} - D_i)].$$

Bidder k would not bid on job i if $pf_{ik} < 0$. When $pf_{ik} \geq 0$ and bidder k can finish job i within the time window, there will be no penalties. Then the maximum discount

λ'_{ik} that bidder k may offer for job i is $\lambda'_{ik} \leq v_i - \rho_{ik}$. Note that there is a minimum increment of discount offer, $\varepsilon > 0$, for each round of bid. Let N_{ik} be the number of increments that bidder k can make on job i . Then $N_{ik} \leq (v_i - \rho_{ik})/\varepsilon$.

With such an upper bound, a bidder can not bid on a job for infinite times. Whenever no bidder offers new bids, the auctioneer will end the auction. Thus the auction will be ended in finite rounds as stated in the following theorem:

Theorem 1: *The reverse auction solution algorithm described in section 3 terminates in finite rounds.*

As the reverse auction is essentially a non-cooperative game problem, let us consider a game theoretic definition of an equilibrium solution to the problem.

Definition: equilibrium solution

A solution to the reverse auction problem is an equilibrium solution if and only if no single bidder can benefit from unilaterally changing the bidder's decision given other bidders' decisions.

It is easy to see that when the reverse auction terminates, no bidders would change their bids under the terminating assignments and discount offers. Otherwise, there would have been new bids and the auction would not have terminated. So, we have the following conclusion:

Theorem 2: *The reverse auction-based algorithm leads to an equilibrium solution.*

4.2 Optimality of bidder's LR-based scheduling algorithm

This numerical study illustrates, via two simple test cases, that the LR-based solution algorithm described in section 3.1 achieves near-optimal job-to-bid selection and schedule under different loading intensity. Let L^D be the cost of dual solution to (D) and J be the cost of the solution to the primal problem (BDP/D) obtained from feasibility adjustment of the dual solution. Note that $L^D \leq J$. The optimality metric chosen is duality gap $\equiv (J - L^D)/|J| \geq 0$; the smaller the gap, the closer the solution to the optimum.

1. Medium load case

The loading intensity is 46.66% of the capacity in this case. Input data and results are given in Appendix A. How the duality gap varies as the LR iteration is shown in figure 3. Since the duality gap approaches zero, an optimal solution is achieved in this case.

2. Overload case

Its loading intensity is 113.33% of the capacity in this case. The scheduling is expected to be difficult and the dual solution will have many capacity constraint violations. Input data, output data and resultant job sequence are given in Appendix B. Figure 4 shows that the duality gap evolves with respect to the LR iteration to 1.89%.

These preliminary results indicate that our LR solution method leads to near-optimum solutions.

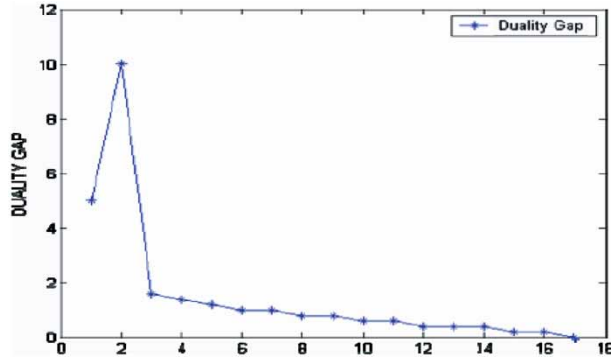


Figure 3. Duality gap of medium loading intensity pattern.

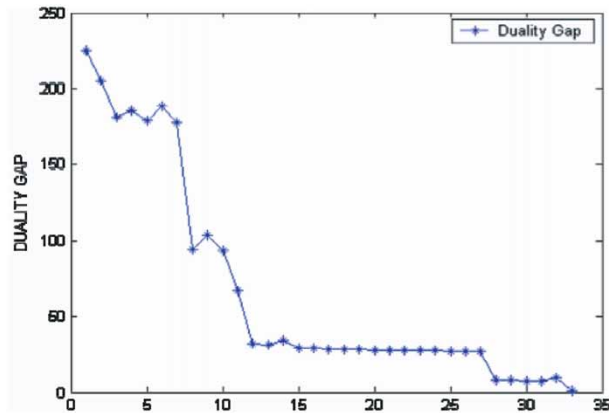


Figure 4. Duality gap of overload loading intensity pattern.

4.3 Effects of formulation and algorithm on bidder's schedule

In the proposed reverse auction, bidders' decisions (BDP) obviously require intensive decision power because of the complexity of fab operations and significantly affect the auction results. The problem formulation and the scheduling algorithm adopted by a bidder reflect the logistic decision capability of the bidder. To assess how individual bidders' logistic decision capability may affect the auction results, this study compares bidder's schedules obtained by using three different combinations of formulation and scheduling algorithm: LR-based algorithm for BDP/D (D-LR), simulation and heuristic-based scheduling algorithm for BDP/D (D-SH), and simulation and heuristic-based scheduling algorithm for BDP/S (S-SH). D-LR represents mean value-based decision with a near-optimal scheduling algorithm while S-SH uses statistical information beyond mean values to make risk-based decision with simple scheduling heuristics. D-SH is the baseline among the three and serves the purpose of a baseline for comparison. In assessing single-bidder decision, what a bidder schedules to bid on are assumed to be what a bidder actually gets assigned. For conciseness of the following discussions, D-LR, S-SH and

D-SH will represent the schedules obtained by using the three combinations, respectively.

The assessments are carried over high and low loading intensity cases, respectively. In each case, there are 100 jobs to be assigned over 400 units of time. The release time of a job is uniformly distributed in $[1, 400]$ while its due time equals the release time plus a time duration randomly generated from $[2, 10]$. Let the time window of a job be $W \equiv \text{due time} - \text{release time}$ and uniform random variables $\alpha, \gamma \in (0, 1)$ and $\beta \in (0, 2)$. The maximum payment to a job, v_i , is generated uniformly in $[1, \text{integer}(\beta W + 3)]$ and mean processing time of a job, μ_i , in $[1, \text{integer}(\alpha W + 1)]$, and the operation cost, ρ_i , in $[1, \text{integer}(\gamma \mu + 1)]$, where $\text{integer}(x)$ takes the integral part of a real number x . Earliness and minimum tardiness penalty coefficients are random in 1 or 2 and the tardiness penalty coefficient is random integer from 1 to 3. Job processing times are exponentially distributed.

Thirty simulation replications are run for each of D-LR, D-SH and S-SH to evaluate their respective bidder's expected profits according to the objective function defined by equation (13). It is expected that D-LR, D-SH and S-SH have significant performance differences under high loading intensity and that D-SH has the poorest performance. Results are as follows.

1. High load case

Under a loading intensity of 86.75%, S-SH uniformly outperforms the other two both across the replications as shown in figure 5 and in terms of profit mean and variance as listed in table 2. Table 2 shows that D-LR is optimistic in bidding for jobs in lack of risk consideration and has the highest number of jobs (84) while S-SH is relatively conservative and has the lowest (73) among the three schedules. In depth analysis of simulation results of D-LR indicates that the LR-based near-optimal scheduling algorithm generates a very tight schedule, where there is little buffering time among jobs. When evaluated by simulation, the processing time outcome of one job may be much longer than the mean value used for D-LR scheduling. Under a tight schedule, the delay of one job may cause delays in many of the subsequent jobs and leads to high over-due penalties and low total job processing profits. Such a phenomenon happens in six replications (1, 2, 7, 9, 28 and 30, figure 5).

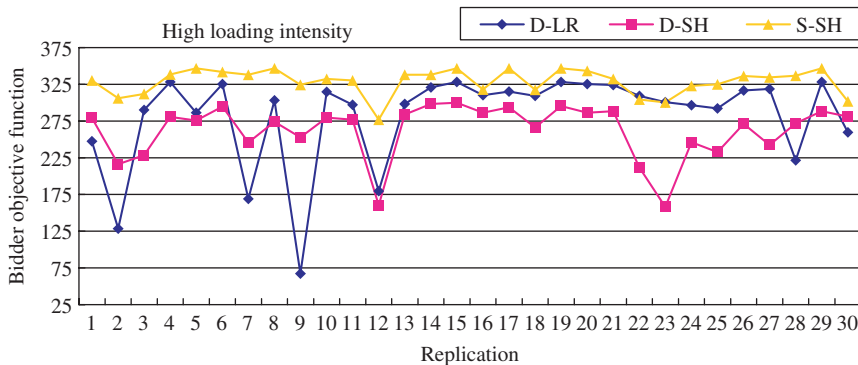


Figure 5. Bidder profits over replications (high loading intensity case).

Table 2. Bidder statistics (high loading intensity case).

	Average bidder profit	Standard deviation of bidder profit	No. of jobs bid	% profit difference with D-SH*
D-SH	262.2	37.1	80	0
D-LR	281.3	65.0	84	7.8%
S-SH	328.7	17.4	73	25.4%

*% difference of A vs. B = (value of A – value of B)/value of B*100.

Table 3. Bidder statistics (low loading intensity case).

	Average bidder profit	Standard deviation of bidder profit	No. of jobs bid	% profit difference with D-SH*
D-SH	398.1	11.8	95	0
D-LR	407.2	5.2	95	2.3%
S-SH	395.9	3.7	93	-0.5%

*% difference of A with B = (value of A – value of B)/value of B*100.

As hypothesized, D-SH indeed has the poorest profit value in all the replications but the aforementioned six, where the use of heuristic gives more time for buffering the processing time uncertainty. Namely, D-SH is more robust to processing time uncertainties than D-LR. This is also supported by the smaller variance of D-SH than that of D-LR. Although S-SH has the least number of jobs, it is uniformly the best in profit over all replications. Such a result indicates the value of using statistical information and risk consideration. The capture of uncertainties in the stochastic problem formulation in conjunction with simple heuristic solution may lead to a schedule superior to the schedule obtained by a near-optimal solution method over a mean value-based problem formulation.

2. Low load case

When loading intensity is low, 43.375% of the capacity in this case, the relative performance among D-LR, D-SH and S-SH is different from that in the high loading intensity case. Table 3 shows that D-LR, D-SH and S-SH bid on 95, 95, and 93 jobs out of the 100 jobs, respectively. This is of no surprise when loading intensity is low. Although S-SH is the most conservative in job bidding and has the lowest average profit, it has the lowest variance of profit. Figure 6 shows that D-LR is superior to the other two in most of the replications and has a low variance too. The average profit of D-LR is slightly (about 3%) higher than those of D-SH and S-SH. Such superiority may be due to the use of the near-optimal LR algorithm. The average profit of D-SH is about the same as that of S-SH. In both the high and low intensity cases, D-SH obvious has a high profit variance than S-SH. Such a result indicates the value of capturing the processing time uncertainty in the problem formulation when SH is used as the solution method.

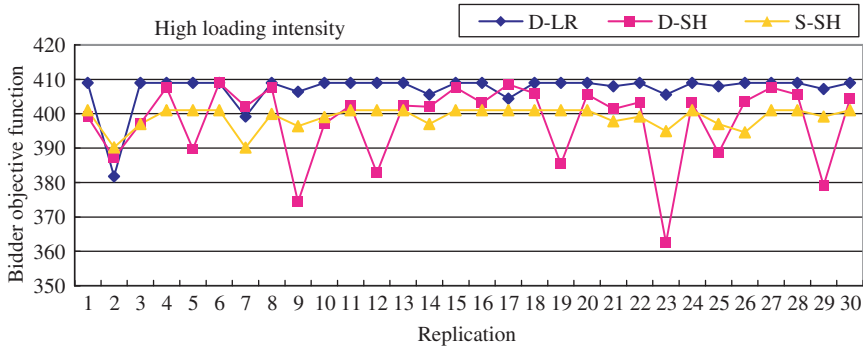


Figure 6. Bidder profits over replications (low loading intensity case).

4.4 Effects of formulation and algorithm on auctioneer's assignment

This experiment evaluates how bidders' problem formulation and solution algorithm affect the job assignment and the auctioneer's cost. There are two bidders, who have the same capacity and 100 jobs to bid over a period of 400 time units. Job parameters are generated in the same way as that of section 4.3 except that the loading intensity adjustment parameter $\alpha\bar{y}$ is doubled in value. The notations D-LR, D-SH, and S-SH now represent the auction decisions for the three different combinations of bidders' problem formulation and scheduling algorithm respectively.

Again, the two cases of low and high loading intensity are investigated, respectively. Key performance metrics include the number of jobs assigned, auctioneer's total cost and the job winning discounts offered by the bidders. In the reverse auction, as the bidders have to take all the risks due to their respective uncertainties, it is natural to hypothesize that S-SH will result in the least numbers of assigned jobs and the lowest bidders discount offers among D-LR, D-SH, and S-SH. It is also expected that bidders' winning discount offers in the low loading intensity case will be higher than those of the high intensity case because there is less job competition in the former case.

Simulation results in figures 7 and 8 depict bidders' conservatism under S-SH in both cases, where the winning discount offers and the numbers of bided jobs are the lowest among the three combinations. Statistics in tables 4 and 5 show that average discounts in the low loading intensity case are higher than those in the high loading intensity case. Such observations are consistent with what are hypothesized and suggest that when loading intensity is low, the three combinations are about equally good and that when loading intensity is high, S-SH leads to a superior result for the bidders. We further reason that since the bidders take all the risk due to uncertainty, the auctioneer may raise the maximum payment values to induce more job bidding from the bidders.

5. Conclusions

Motivated by foundry service provisioning and the emergence of foundry brokerage in the semiconductor industry, this paper adopted a reverse auction-based

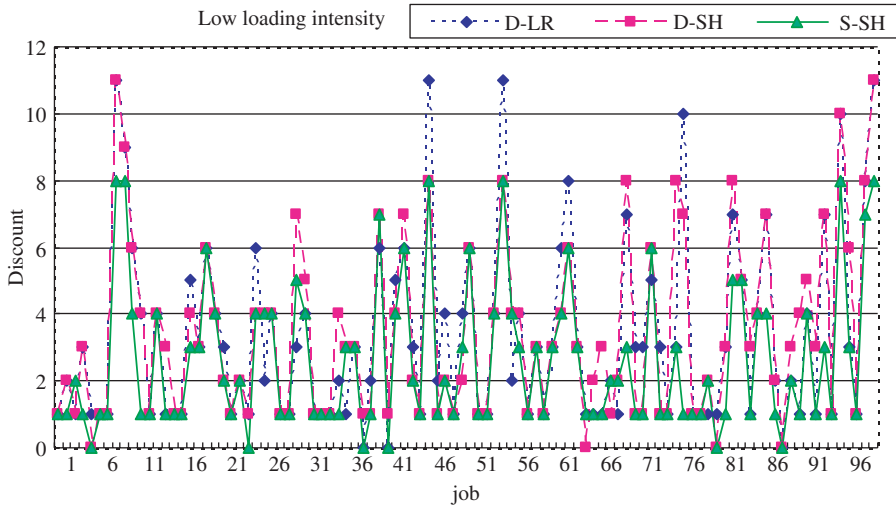


Figure 7. Discount for each job under low loading intensity.

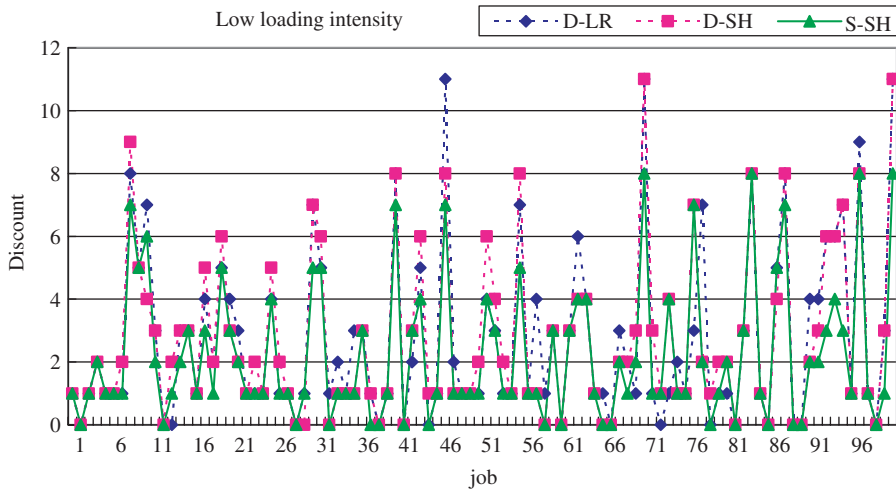


Figure 8. Discount for each job under high loading intensity.

Table 4. Auction statistics (low loading intensity).

	Auctioneer's cost		Number of unassigned jobs	Winning discount (%)	
	Average	Difference with D-SH (%)		*Average	Std. dev.
D-SH	481	0	5	37.08	18.32
D-LR	491	2.1	3	37.36	19.87
S-SH	557	16.9	6	29.90	16.89

*%Winning discount = (job discount/job value)*100.

Table 5. Auction statistics (high loading intensity).

	Auctioneer's cost		Number of unassigned jobs	Winning discount (%)	
	Average	Difference with D-SH (%)		*Average	Std. dev.
D-SH	570	0	16	30.03	21.73
D-LR	565	0.9	14	29.58	21.54
S-SH	617	8.2	18	24.18	17.76

*%Winning discount = (job discount/job value)*100.

mechanism to model job order assignment between a job owner and a few qualified fabs. With the simplifying assumptions of single-operation jobs and no job transfers or capacity backups among fabs, the model focuses on the non-cooperative gaming because of private information among the job owner and the fabs such as objectives, valuation of jobs, available capacity and constraints. In addition to a mathematical formulation of the reverse auction problem for job assignment, solution algorithms for both bidder and auctioneer are proposed to model the job assignment decision of the job order owner and the bid selection and production scheduling decision of the fabs. A simple analysis has shown that the auction algorithm converges to an equilibrium, where no single bidder would unilaterally deviate from the auction result. Numerical experiments with the reversed auction model demonstrate that consideration of uncertainties in bidder's decision has a larger impact on performance of both bidders and the auctioneer than optimality of the mean value-based scheduling algorithm that a bidder adopts.

Acknowledgements

This work was supported in part by the Semiconductor Research Corporation and International Semiconductor Manufacturing Initiative under FORCE-II project 1214 and by the National Science Council, Taiwan, R.O.C., under grants NSC-93-2213-E-002-043 and 94-2213-E-002-015.

Appendix A

Data of medium loading intensity pattern

	Release date	Due date	Earliness/ tardiness penalty	Step penalty	Value	Operation cost	Process time
Job 1	2	6	1/1	1	8	1	2
Job 2	5	7	1/2	1	10	2	3
Job 3	9	11	2/2	1	6	1	2
Job 4	13	15	1/1	1	5	1	2
Job 5	13	16	1/1	1	6	1	3
Job 6	10	14	2/2	1	7	2	2

	Release date	Due date	Earliness/ tardiness penalty	Step penalty	Value	Operation cost	Process time
Job sequence of medium loading intensity pattern							
Sequence	1	2	3	4	5	6	
Job	1	2	3	6	4	5	

Appendix B

	Release date	Due date	Earliness/ tardiness penalty	Step penalty	Value	Operation cost	Process time								
Job 1	2	6	1/2	1	8	1	2								
Job 2	19	22	1/1	1	6	2	3								
Job 3	3	5	2/3	1	6	1	4								
Job 4	9	10	2/2	1	5	1	1								
Job 5	3	4	3/1	1	3	1	1								
Job 6	15	16	2/2	1	7	2	2								
Job 7	1	4	1/1	1	4	1	1								
Job 8	3	9	2/2	1	6	1	1								
Job 9	16	18	1/1	1	9	2	3								
Job 10	2	6	2/1	1	8	2	1								
Job 11	20	22	1/2	1	4	1	4								
Job 12	13	14	3/3	1	6	2	1								
Job 13	15	16	1/1	1	9	1	3								
Job 14	15	20	4/2	1	7	1	2								
Job 15	11	13	3/1	1	4	1	5								
Job sequence of overload pattern															
Sequence	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Job	7	10	5	1	8	4	13	12	6	9	14	2	0	0	0

References

- Bertsekas, D.P., The auction algorithm for assignment and other network flow problems: a tutorial. *Interfaces*, 1990, **20**, 133–149.
- Clarke, P., Foundry broker sets up shop with 1st Silicon, Tower, ZMD. *EE Times*, March 2004. Available online at: <http://www.eetimes.com/news/latest/showArticle.jhtml?articleID=18400118>.
- Dewan, P. and Joshi, S., Auction-based distributed scheduling in a dynamic job shop environment. *Int. J. Prod. Res.*, 2002, **40**(5), 1173–1191.
- Fan, M., Stallaert, J. and Whinston, A.B., Decentralized mechanism design for supply chain organizations using an auction market. *Inform. Syst. Res.*, 2003, **14**(1), 1–22.
- Ghawi, D. and Schneider, G.P., New Approaches to Online Procurement, in *2004 Proceedings of the Academy of Information and Management Sciences*, Vol. 8, Number 2, Allied Academies International Conference, Maui, 13–16 October 2004, pp. 25–28.
- Hao, S., A study of basic bidding strategy in clearing pricing auctions. *IEEE Trans. Power Syst.*, 2000, **15**(3), 975–980.

- Hogan, W.E., Market design and electricity restructuring, in *Association of Power Exchanges 2005 Annual Conference*, Orlando FL, Presentation, 1 November 2005.
- Hoitomt, D.J., Luh, P.B. and Pattipati, K.R., A practical approach to job-shop scheduling problems. *IEEE Trans. Robot. Autom.*, 1993, **9**(1), 1–13.
- Hwang, T.K. and Chang, S.C., Design of a Lagrangian relaxation-based hierarchical production scheduling environment for semiconductor wafer fabrication. *IEEE Trans. Robot. Autom.*, 2003, **19**(4), 566–578.
- Karabuk, S. and Wu, S.D., Decentralizing semiconductor capacity planning via internal market coordination. *IIE Trans.*, 2002, 743–759.
- Karabuk, S. and Wu, S.D., Incentive schemes for semiconductor capacity-allocation: a game theoretic analysis, 2005, **14**(2), 175–188.
- Kutanoglu, E. and Wu, S.D., On combinatorial auction and lagrangean relaxation for distributed resource scheduling. *IIE Trans.*, 1999, **31**, 813–826.
- Liao, D.-Y., Chang, S.C., Pei, K.W. and Chang, C.M., Daily scheduling for R&D semiconductor fabrication. *IEEE Trans. Semiconductor Manuf.*, 1996, **9**(4), 550–561.
- Luh, P.B., Ni, M., Chen, H. and Thakur, L.S., Price-based approach for activity coordination in a supply network. *IEEE Trans. Robot. Autom.*, 2003, **19**(2), 335–346.
- Luh, P.B., Blankson, W.E., Chen, Y., Yan, J.H., Stern, G.A., Chang, S.C. and Zhao, F., Payment cost minimization auction for deregulated electricity markets using surrogate optimization. *IEEE Trans. Power Syst.*, 2006, **21**(2), 568–578.
- Milgrom, P.R. and Weber, R.J., A theory of auctions and competitive biddings. *Econometrica*, 1982, **50**(5), 1089–1122.
- Sarne, D., Hadad, M. and Kraus, S., Auction equilibrium strategies for task allocation in uncertain environments. In *Cooperative Information Agents VIII, LNAI*, edited by S. Ossowski and V. Kashyap, pp. 271–285 (Springer-Verlag).
- Srivinas, M., Tiwari, K. and Allada, V., Solving the machine-loading problem in a flexible manufacturing system using a combinatorial auction-based approach. *Int. J. Prod. Res.*, 2004, **42**(9), 1879–1893.
- Su, Y.-H., Guo, R.S. and Chang, S.C., Virtual fab: an enabling framework and dynamic manufacturing service provision mechanism. *Inform. & Manag.*, 2005, **42**(2), 329–348.
- Tang, L., Luh, P.B., Liu, J. and Fang, L., Steelmaking process scheduling using Lagrangian relaxation. *Int. J. Prod. Res.*, 2002, **40**(1), 55–70.
- de Vries, S. and Vohray, R.V., Combinatorial auctions: a survey. *INFORMS J. Comput.*, 2003, **15**(3), 284–309.
- Yu, D., Luh, P.B. and Chang, S.C., Achieving six sigma deliveries in supply chains, in *Proceedings of the 16th IFAC World Congress*, Prague, 2005, 4–8 July.
- Zhao, X., Luh, P.B. and Wang, J., The surrogate gradient algorithm for Lagrangian relaxation method. *J. Opt. Theory Appl.*, 1999, **100**(3), 699–712.