

주문형반도체 제조회사의 작업하달 논리의 개발

김기영^{1*} · 강창호² · 김갑환²

¹동서대학교 국제관계학부 국제물류전공 / ²부산대학교 공과대학 산업공학과

An Order Releasing Algorithm for a Semiconductor Wafer Manufacturer

Ki Young Kim¹ · Chang Ho Kang² · Kap Hwan Kim²

¹Department of International Logistics, Division of International Relations, Dongseo University

²Department of Industrial Engineering, Pusan National University

In this paper, based on the process analysis of a semiconductor company, a lot order release procedure (input regulation) is developed for a semiconductor company. The major characteristic of the order release procedure in this paper is to consider the workloads of machines which are obtained from a virtual lot flows by the dispatching rule at machines in the shop of the semiconductor company. The objective is to minimize the cycle time and to obtain other good performances. A simulation is performed in order to evaluate the order release procedure in this paper.

Keyword: Order release, Input regulation, Semiconductor, Simulation, WIP, Dispatching

1. 서론

본 연구의 목표는 다양한 주문형 반도체를 생산하는 회사를 대상으로 하여 해당업체가 본격적인 생산단계에 이르렀을 때에 작업장의 공정재고를 최적으로 유지하기 위한 작업하달 논리를 개발하는 것이다.

대상업체는 주문형 반도체를 생산하기 때문에 고객주문에 따라 웨이퍼의 가공경로가 다양할 뿐만 아니라 가공공정의 수가 많고 동일한 가공기계를 재방문해서 가공되는 경우도 많다. 따라서 작업장의 공정재고를 합리적인 수준으로 유지하기 위하여 본 논문에서는 작업장의 웨이퍼들을 담는 상자를 단위로 하는 로트의 투입시점, 투입할 로트의 선정에 대해서 다룬다.

반도체와 관련되면서 본 연구의 내용과 관련된 연구들로는 다음과 같은 것들이 있다. Sung and Choung(2000)은 반도체 제조공정중 배치로 가공되는 오븐공정에 대해서 배치가공시간을 최소로 하는 문제에 대해서 연구하였다. Sivakumar and

Chin(2001)은 시뮬레이션 모델을 사용하여 반도체 후속공정들에 있어서 사이클타임과 수율을 분석하였다. Lee et al.(1997)은 변동하는 사이클타임을 고려하여 반도체 웨이퍼 제조회사에서의 생산계획에 대해서 연구하였다. 이들의 연구에서는 실제 웨이퍼제조에서 사이클타임의 분산이 크고 또한 공정상에서의 변동에 의해 사이클타임의 분포가 한쪽으로 치우쳐져 있다는 점을 감안하여 변동하는 사이클타임과 작업장 상태를 함께 고려하여 생산계획에 반영하였다. Dabbas et al.(2001)은 반도체 제조시스템에서 복수의 디스패칭규칙을 혼합하여 하나의 디스패칭규칙을 만들고 이를 적용하여 복수의 목적함수를 동시에 최적화하려고 시도하였다. Vargas-Villamil and Rivera(2000)는 반도체 제조라인의 특성의 하나인 이산사건 복귀라인(discrete event reentrant lines)에 대해서 두단계의 계층구조로 스케줄링하였다. Kang(1996)은 이벤트 기반의 최적화와 시뮬레이션을 기반으로 반도체 웨이퍼 제조업체에 대한 스케줄링 방법에 대해 연구하였다. Lin(1996)은 피드백정보와 예측

*연락처 : 김기영, 617-716 부산광역시 사상구 주례2동 산69-1번지 동서대학교 국제관계학부 국제물류전공, Fax : 051-320-1630,

E-mail : kiykim@gdsu.dongseo.ac.kr

2005년 1월 접수, 2회 수정 후 2006년 3월 게재확정

정보를 이용하여 반도체 웨이프 제조업체에 대한 스케줄링 방법에 대해 연구하였다. Shen(1997)은 추계이론을 이용하여 반도체 웨이프 제조업체에 대한 스케줄링 방법에 대해 연구하였다. Sloan(1998)는 라인상의 장비조건과 수율(yield) 정보를 이용하여 반도체 웨이프 제조업체에 스케줄링 방법에 대해 연구하였다.

참고로, 본 연구의 내용과 관련하여 일반적인 제조업체를 대상으로 한 연구들을 논문초록 수준에서 검토한 것으로는 다음과 같은 것들이 있다. Koh and Saad(2004)는 MRP로 관리되는 불확실한 제조시스템의 작업하달을 위해서 리드타임, 능력, 안전재고, 그리고 배치의 유연성을 파라미터로 하는 지능형 피드백에 대해서 연구하였다. Kingsman(2000)은 생산시스템에서 동적 능력계획을 위한 작업부하에 대한 수리모형을 제안하였다. Oosterman et al.(2000)은 시뮬레이션 연구를 통하여 공정의 특징이 작업부하 관리에 미치는 영향을 연구하였다. Donselaar and Dubbels(2002)는 시뮬레이션을 통하여 시스템 재고와 과민성을 최소화하기 위한 작업하달 방안을 제안하였다. Land and Gaalman (1996)은 작업부하 관리를 위한 기존연구들의 개념들을 비교 연구하여 소개하였다. Ooijen and Bertrand(2001)는 라루팅과 작업부하에 종속된 체류시간의 분포함수에 기반하여 개별생산(job shop)에 있어서 경제적인 납기설정 방법을 제안하고 있다. Corsten et al.(2005)은 유연제조시스템을 위한 작업하달 접근법을 관련된 성격에 기준을 두고 기존연구들을 잘 요약하고 있다. Raaymakers et al.(2001)은 배치가공 제조업체에 있어서 완성시간(makespan)의 추정에 관한 연구하였다. Weng(1996)은 주문생산시스템에서 제조 리드타임 계획문제를 위한 모형을 제안하였다. 이상의 관련 연구들은 사용가능한 학술 데이터베이스 하에

서 조사된 것이므로 본 연구와 관련된 좋은 논문들을 조사하지 못했을 가능성이 크다. 그러나 전반적으로 본 논문의 내용과 기여도를 이해하는데 큰 어려움이 없을 것으로 생각된다.

작업하달의 제어규칙에 대한 기존의 작업하달 연구 중 직접적으로 본 연구와 관련되면서 중요한 것이라 생각되는 것들을 Lin(1996)과 Kim et. al.(1996)의 논문의 내용을 참고하면서 요약 정리한 것이 <표 1>이다.

2. 대상업체의 운영규칙

먼저 대상업체의 운영규칙에 대해서 언급하기 전에 주문상태의 변화에 대해서 살펴 본다. 고객의 주문이 들어오면 주수 및 기술부서에서 기술적 처리과정을 거친다. 여기서는 고객주문은 작업이 가능한 로트로 분할된다. 일단, 로트가 형성되면 그 로트의 속성인 웨이프의 수, 가공경로, 납기일 등의 속성이 할당되고 작업하달 대기열로 이동하게 된다. 작업하달 대기열로 이동된 로트는 작업장에 투입에 대한 작업하달이 있을 때까지 대기한다. 본 연구에서의 로트의 작업하달 시점은 8시간마다 이루어진다고 가정한다. 작업하달되는 대상 로트는 기본적으로 여유시간이 적은 순으로 선정된다.

일단, 작업하달된 로트는 작업장에 투입되어 필요한 공정들을 거치게 된다. 임의의 시점에 작업장에서 머물고 있는 로트의 상태는 크게 가공을 위해 기계 대기열에서 대기중인 경우와 가공중인 경우로 나뉜다. 각 기계에서의 디스패칭 규칙은 여유시간이 적은 순이다. 로트는 모든 공정을 완료하고 나면 최종적으로 고객에게 납품된다.

표 1. 작업하달(order release)의 제어규칙

제 어 규 칩	설 명
PA (Possion Arrival)	Poission분포로 로트를 투입한다.
FA (Fixed inter-arrival time)	일정시간간격으로 로트를 투입한다.
CONWIP (Fixed-WIP control)	작업장에 규정된 공정재고의 수를 유지할 수 있도록 로트를 투입한다.
BWL (Bottleneck workload) by Wein(1988)	병목장비에서의 총부하가 규정된 목표부하 이하로 떨어졌을 때 로트를 투입한다. 단점으로는 목표 부하를 정의하는데 어려움이 있다.
SA (Starvation Avoidance) by Glassey and Resende(1988)	병목장비에서 기계유휴(starvation)가 예상되는 미래의 특정시점에 새로 투입된 로트가 그 장비에 도착하여 기계유휴를 방지할 수 있을 때 현재 새로운 로트를 투입한다. 따라서 새 로트가 병목장비까지 도달할 때까지의 모든 공정에서의 가상재고가 정의되어야 한다.
QM (Queue Management) by Leachman et. al.(1990)	현재의 공정재고 위치, 장비처리능력, 시간에 따른 장비로의 로트의 예상도착시간을 예측하기 위해서 평균체류시간(mean flow time)의 추정값을 이용한다. 미래시점에서 모든 장비에서 안전수준 이하로 부하가 떨어지는 것이 예상될 때 로트를 투입한다.
DEC (DEscending Control) by Glassey et. al.(1993)	state space로서 재고벡터를 정의하고 state의 subset로 제어집합(control set)을 정의하여 주어진 재고비용 제약하에서 이익을 최대화하는 관점에서 로트를 투입한다.
LR (Lateness release) by Lin(1996)	작업장의 혼잡도의 측도로서 시스템 내에 있는 모든 WIP의 평균 lateness을 사용한다. 따라서 추정된 평균 lateness에 바탕을 두고 로트를 작업장에 투입한다.
SA based projected Queue by Kim et. al.(1996)	특정 미래시점까지 장비에서 기계유휴가 발생하지 않도록 작업장내 한 장비에서라도 안전수준(safety level)이하로 예상대기(projected queue)가 떨어질 경우에 로트를 투입한다. 한 장비에서라도 특정 미래시점까지의 모든 시점에서 예상대기가 안전수준을 초과하면 로트를 투입하지 않는다.

2.1 여유시간을 사용한 디스패칭의 구현

현장 실무자의 의견을 반영하여 도입한 주된 가정은 다음과 같다.

- (1) 디스패칭 규칙으로 최소여유시간 규칙을 사용한다.
- (2) 가공시간은 웨이퍼의 개수에 비례하며 사람이나 기계에 의해 발생하는 변동은 무시한다.
- (3) 웨이퍼가 가공을 마치고 후속공정으로 이동할 때 걸리는 운반시간은 무시한다.
- (4) 장비의 고장과 웨이퍼의 불량은 발생하지 않는다.
- (5) 대상업체는 3교대로 24시간 모두 가동되는 것으로 가정한다.
- (6) 작업장의 각 기계 대기열에서 로트들이 대기를 마칠 때, 해당 대기열에서의 여유시간과 대기시간에 대한 각각의 평균과 분산은 갱신된다.

먼저 본 논문의 내용을 설명하기 위해서 다음과 같은 기호를 정의한다.

- i = 로트의 색인, $i=0, 1, 2, \dots, n$
- j = 기계 혹은 기계 대기열의 색인, $j=0, 1, 2, \dots, m$
- k_i = 로트 i 의 공정의 수
- K = 가공공정의 색인, $k=0, 1, 2, \dots, k_i$
- $b(i,k)$ = 로트 i 의 k 번째의 가공공정을 위한 대기열의 색인, $k=0, 1, 2, \dots, k_i$
- $m(i,k)$ = 로트 i 의 k 번째의 가공공정을 위한 가공기계, $k=0, 1, 2, \dots, k_i$
- $p_{m(i,k)}$ = 로트 i 의 k 번째 공정에서의 가공시간
- $W_{b(i,k)}$ = 대기장소 b 에서 로트의 대기시간
- $S_{b(i,k)}$ = 대기장소 $b(i,k)$ 에서 로트의 여유시간
- $\tilde{W}_{b(i,k)}$ = 대기장소 $b(i,k)$ 에서 대기시간의 평활화된 평균
- $\tilde{S}_{b(i,k)}$ = 대기장소 $b(i,k)$ 에서 여유시간의 평활화된 평균
- $\tilde{V}(W_{b(i,k)})$ = 대기장소 $b(i,k)$ 에서 대기시간의 분산의 평활화된 평균
- $\tilde{V}(S_{b(i,k)})$ = 대기장소 $b(i,k)$ 에서 여유시간의 분산의 평활화된 평균
- $A_{b(i,k)}$ = 로트 i 의 k 번째 공정을 위해서 대기장소 $b(i,k)$ 에 도착하는 시각
- d_i = 고객의 주문에 상응하는 로트 i 의 납기일

각 로트들이 기계 대기장소 $b(i,k)$ 에서 떠날 때 여유시간과 대기시간에 대한 각각의 평균과 분산에 대한 추정치는 다음과 같이 갱신된다고 가정한다.

$$\begin{aligned} \tilde{S}_{b(i,k)} &= (1 - \alpha_s) \tilde{S}_{b(i,k)} + \alpha_s S_{b(i,k)} & (1) \\ \tilde{W}_{b(i,k)} &= (1 - \alpha_w) \tilde{W}_{b(i,k)} + \alpha_w W_{b(i,k)} & (2) \\ \tilde{V}(S_{b(i,k)}) &= (1 - \beta_s) \tilde{V}(S_{b(i,k)}) + \beta_s (S_{b(i,k)} - \tilde{S}_{b(i,k)})^2 & (3) \end{aligned}$$

$$\tilde{V}(W_{b(i,k)}) = (1 - \beta_w) \tilde{V}(W_{b(i,k)}) + \beta_w (W_{b(i,k)} - \tilde{W}_{b(i,k)})^2 \quad (4)$$

단, $\alpha_s, \alpha_w, \beta_s,$ and β_w 는 0초과 1미만의 값이다.

작업하달을 하기 위해서는 작업하달될 로트들에 의해 앞으로 각 기계에서 걸리게 될 부하를 고려하여야 한다. 이를 위해 우선, 작업장에 이미 작업하달되어 있는 로트들과 작업하달될 로트들이 각 기계대기장소에 도착하게 되는 미래시점을 계산하여야 한다. 그런데, 디스패칭 규칙으로 최소여유시간 규칙이 사용되므로 각 기계 대기장소에서의 대기시간을 추정하기가 곤란하고 따라서 각 기계에 도착하게 되는 시점을 계산하기가 어렵다. 이를 극복하기 위해 본 논문에서는 기계 대기장소 $b(i,k)$ 에서 대기중인 어떤 로트의 여유시간의 표준화된 값과 대기시간의 표준화된 값은 동일하다는 가정을 도입한다. 이는 기계 대기열들에서 경험하는 어떤 로트의 여유시간이 그 평균 여유시간에서 특정 값으로 벗어날 확률과 해당 로트가 해당 대기열에서 경험하는 대기시간이 그 평균대기시간에서 특정 값으로 벗어날 확률은 동일하다고 가정하는 것과 같다.

$$W_{b(i,k)} = \tilde{W}_{b(i,k)} + \frac{S_{b(i,k)} - \tilde{S}_{b(i,k)}}{\sqrt{\tilde{V}(S_{b(i,k)})}} \sqrt{\tilde{V}(W_{b(i,k)})} \quad (5)$$

식 (5)를 이용해서 로트 i 의 k 번째 공정을 위해 기계 $m(i,k)$ 의 대기열 $b(i,k)$ 에 대기중인 로트 i 의 여유시간이 $S_{b(i,k)}$ 는 다음과 같이 표현할 수 있다.

$$S_{b(i,k)} = d_i - \tau - R_{ik} \quad (6)$$

단, R_{ik} 는 k 번째 공정을 위해 기계 대기장소에서 대기중인 로트 i 의 잔여체류시간이다.

잔여체류시간 R_{ik} 는 모든 후속공정에서의 대기시간과 가공시간의 합이다. 결국 식 (6)의 여유시간을 평가하기 위해서는 모든 후속공정의 대기장소에서의 대기시간을 알아야 한다는 의미인데, 평가시점에서는 이를 알 수 없고 또한 운영규칙에 의존적이다. 즉, $S_{b(i,k)}$ 와 R_{ik} 는 서로 종속적이다. 따라서 $S_{b(i,k)}$ 의 값을 추정하기 위해서 본 논문에서는 기계 $m(i,k)$ 의 디스패칭 시점에 로트 i 의 여유시간을 구하는 다음과 같은 절차를 도입한다.

여유시간을 추정하는 절차

편의상 $m(i,k)=j$ 라 두자. 그리고 기계 대기열 $b(i,k)$ 에서 로트 i 의 대기시간 추정치를 $W_{b(i,k)}$ 라 두자.

(단계 0): (초기화) 기계 $m(i,k)$ 에서 대기중인 로트 i 가 이후 거치게 될 기계들에 대해서 $W_{b(i,k)} = \tilde{W}_{b(i,k)}$ 라 둔다. 단, $k=1, \dots, k_i$.

(단계 1): (잔여가공시간의 추정) 기계 j 에서 로트 i 의 잔여가공시간과 잔여대기시간에 대한 추정치를 각각 r_{ij} 와 w_{ij} 라

두자. 기계 $j=m(i,k)$ 에서 가공중인 로트의 잔여가공 시간을 r_{ij} 로 두고, $w_{ij} = \max\{0, \tilde{W}_{b(i,k)} - w_i\}$ 로 둔다. 단, w_i 는 기계 j 에서 로트 i 의 현재까지의 실제 대기시간을 나타낸다. 그렇지 않으면 $r_{ij} = p_{ik}$ 그리고 $w_{ij} = 0$ 라 둔다. $s = k$.

(단계 2): (예상 여유시간과 대기시간의 계산) $s = s + 1$. 만약 $s > k_i$ 이면 단계 3으로 간다.

$$S_{b(i,s)} = d_i - \tau - r_{ij} - w_{ij} - \sum_{v=s+1}^{k_i} (p_{m(i,v)} + W_{b(i,v)}).$$

$$W_{b(i,s)} = \tilde{W}_{b(i,s)} + \frac{S_{b(i,s)} - \tilde{S}_{b(i,s)}}{\sqrt{V(S_{b(i,s)})}} \sqrt{V(W_{b(i,s)})}$$

$$A_{b(i,s)} = \tau + r_{ij} + w_{ij} + \sum_{v=s+1}^{k_i} (p_{m(i,v)} + W_{b(i,v)})$$

만약, $s > k_i$ 이면 단계 3으로 간다. 그렇지 않으면 단계 2로 간다.

(단계 3): (여유시간의 추정)

$$S_{b(i,k)} = d_i - \tau - r_{ij} - w_{ij} - \sum_{v=k+1}^{k_i} (p_{m(i,v)} + W_{b(i,v)})$$

여유시간을 추정하는 절차의 예

<그림 1>과 같은 작업장을 가정하자. 어떤 임의의 시점에 M-A에서의 가공중인 어떤 로트의 작업이 완성되어 기계 M-A가 가용하게 되었다고 하고 이 시점을 0이라 가정하자. 이 경우에 M-A의 앞에서 대기하고 있는 로트 1, 2, 3 중에서 어느 로트를 먼저 작업할 것인가를 결정해야 한다. 이를 위해 로트 1, 2, 3의 각각에 대한 여유시간을 모두 알아야 한다. 여기서는 편의상 향후 기계 M-A, M-B, M-C, M-D, M-E의 순으로 거칠 로트 2에 대해서만 여유시간을 구하는 예를 제시하겠다. 로트 2의 납기일은 180일이라 가정하고, 편의상 로트2의 M-A에서의 시점 0까지의 대기시간은 0이라 가정하자.

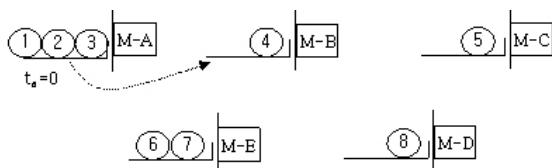


그림 1. M-A에서 가공할 로트를 선정하는 시점에서의 작업장 상황.

표 2. 로트 2가 거칠 기계에서의 대기시간과 여유시간의 각각의 평균과 분산

기 계	M-A	M-B	M-C	M-D	M-E	완료시점
가공시간	10	10	10	10	10	170
대기시간 E[W]	20	10	30	20	40	
대기시간 $\sqrt{V[W]}$	2	3	2	3	2	50
여유시간 E[S]	10	10	5	10	15	
여유시간 $\sqrt{V[S]}$	3	2	1	2	4	

<표 2>는 기계 M-A에서 로트 1, 2, 3의 여유시간을 계산하는

시점에서 향후 로트 1, 2, 3이 거치게 되는 기계들에 대해 각 기계에서 지금까지 갱신하면서 유지해 온 여유시간과 대기시간에 대한 각각의 평균과 분산을 읽어 온 것이라 가정하자.

이상과 같은 조건에서 로트2의 후속공정에서의 예상 대기시간과 예상 여유시간을 구해 보자. 만약 로트 2가 M-A에 바로 투입된다고 가정하면, M-A의 가공시간 만큼 경과한 시간이 M-B에서는 로트 2의 도착시각이 된다. M-B에서 로트 2의 예상 여유시간과 예상 대기시간은 각각 다음과 같다.

$$S_{b(2,2)} = d_2 - \tau - R_{22} = 180 - 10 - \{(10 + 30 + 20 + 40) + (10 + 10 + 10 + 10)\} = 30$$

$$W_{b(2,2)} = E[W_{b(2,2)}] + \frac{S_{b(2,2)} - E[S_{b(2,2)}]}{\sqrt{V[S_{b(2,2)}]}} \sqrt{V[W_{b(2,2)}]} = 10 + \frac{30 - 10}{2} \cdot 3 = 40$$

로트 2의 M-C에 도착시점은 M-A에서의 가공시간 10, 기계 M-B에서의 대기시간 40과 가공시간 10이 소요되어 60이 된다. 따라서 M-B에서 로트2의 예상 여유시간과 예상 대기시간은 각각 다음과 같다.

$$S_{b(2,3)} = d_2 - \tau - R_{23} = 180 - 60 - \{(30 + 20 + 40) + (10 + 10 + 10)\} = 0$$

$$W_{b(2,3)} = E[W_{b(2,3)}] + \frac{S_{b(2,3)} - E[S_{b(2,3)}]}{\sqrt{V[S_{b(2,3)}]}} \sqrt{V[W_{b(2,3)}]} = 30 + \frac{(0 - 5)}{1} \cdot 2 = 20$$

동일한 방식으로 M-D, M-E에서의 예상 여유시간과 예상 대기시간은 각각 다음과 같다.

$$S_{b(2,4)} = 180 - 90 - \{(20 + 40) + (10 + 10)\} = 10$$

$$W_{b(2,4)} = 20 + \frac{(10 - 10)}{2} \cdot 3 = 20$$

$$S_{b(2,5)} = 180 - 120 - \{(40) + (10)\} = 10$$

$$W_{b(2,5)} = 40 + \frac{(10 - 15)}{4} \cdot 2 = 37.5$$

<표 3>은 M-A에서 로트 2가 가공로트로 선정될 때, 로트 2의 후속기계에서의 예상 도착시간, 예상 대기시간, 가공시간, 완료시점을 제시하고 있다.

표 3. 로트 2가 선택될 경우 각 기계에서의 예상 대기시간과 예상 여유시간

	M-A	M-B	M-C	M-D	M-E
예상 기계도착시간	0	10	60	90	120
예상 여유시간		30	0	10	10
예상 대기시간		40	20	20	37.5
가공시간	10	10	10	10	10

결과적으로 M-A에서의 로트 2의 예상 여유시간은 <표 3>을 이용하여 다음과 같이 계산할 수 있다.

$$S_{b(2,1)} = 180 - 0 - \{(40 + 20 + 20 + 37.5) + (10 + 10 + 10 + 10 + 10)\} = 12.5$$

참고로, 이때의 로트 2의 예상 대기시간은 다음과 같다.

$$W_{b(2,1)} = 20 + \frac{12.5 - 10}{3} \cdot 2 = 20 + 1.7 = 21.7$$

동일한 방식으로 로트 1과 3에 대해서도 동일한 방법으로 예상 여유시간을 계산할 수 있다. 일단, M-A기계에서 로트 1, 2, 3의 예상 여유시간을 구하게 되면, 예상 여유시간이 적은 순으로 가공할 로트를 선정할 수 있다. 또한, 로트 1, 2, 3이 후속공정의 기계에서의 도착시간도 추정할 수 있다.

수학적으로 엄밀히 증명할 수는 없었지만 이상의 여유시간 추정을 위한 절차의 주된 장점은 다음과 같다 것들이 있다고 생각된다.

첫째, 예상 여유시간을 활용하여 예상 대기시간을 구하는 방식은 직관과 많은 부분 일치한다.

둘째, 여유시간 추정을 위한 절차에서 추정하는 여유시간은 불편추정치에 가깝다. 기계 대기열 $b(i, k)$ 에서의 예상 여유시간의 추정치와 예상 여유시간의 기대값을 차감하면 다음과 같다.

$$\hat{S}_{b(i,k)} - E[S_{b(i,k)}] = \sum_{v=s+1}^{k_i} (-W_{b(i,v)} + E[W_{b(i,v)}])$$

위의 함수의 오른쪽 항은 $\sum X_i$ 의 꼴이다. X_i 가 임의의 정규분포를 따를 때 적당히 큰 n 의 값에 대해서 $P(\sum X_i = 0) \approx 1$ 이 성립한다. 본 논문에서는 작업하달 규칙으로 일정시간 간격으로 CONWIP을 사용하여 로트를 투입한다고 가정하였다. 따라서 시간이 흐름에 따라 작업장의 작업부하 혹은 각 기계에서의 평균대기시간에 상승추세 혹은 하강추세가 지속될 가능성은 낮기 때문에 여유시간의 추정치와 여유시간의 기대값을 차감한 값은 0의 값에 가까워 질 것이다.

3. 작업하달의 단계별 절차

기존의 작업하달 방식들은 주로 기계가 일감이 없어서 가동을 멈추는 현상이 발생할 가능성을 감소시키기 위하여 작업하달 대기열의 로트를 가상적으로 하나씩 투입하여 작업부하를 증가시키면서 작업장에 미치는 영향 등을 검토하면서 해당 로트의 투입을 결정해가는 방식을 사용한다.

본 연구에서의 작업하달 논리의 주된 생각은 평균적으로 투입되는 작업하달 로트의 수보다 조금 더 많은 로트를 가상적으로 투입하여 각 시점별, 기계별로 작업부하를 계산한 후, 부하가 과다하면 투입된 가상로트에 해당하는 작업하달 대기열

의 실제 로트를 투입하지 않는 방식이다. 작업하달 대기열의 실제로트에 상응하는 가상으로 투입된 로트를 이후부터 편의상 가상로트라 부르겠다.

이 방식의 첫번째 주된 장점으로서는 일단 가상로트를 투입한 후 문제가 없으면 가상로트에 해당하는 실제 로트를 작업장에 투입하는 것으로 절차를 종료할 수 있으므로 작업하달에 소요되는 계산량을 줄일 수 있는 가능성이 크다는 것이다. 그리고 현실적으로도 작업하달 업무의 간소화를 달성할 수 있는 가능성이 크다. 두번째 주된 장점은 각 시점별, 기계별로 작업부하가 걸리는 시점을 보다 정확하게 계산하기 위해서 각 기계의 대기장소에서의 우선순위규칙인 최소여유시간 규칙을 반영하여 시점별, 기계별로 대기시간을 별도로 예측하는 방식을 사용한다. 즉, 동일한 기계, 동일한 공정, 동일한 로트일지라도 로트들이 경험하는 대기시간은 달라질 수 있다.

작업하달의 절차를 실제 시스템으로 구현하기 위해서 세부 단계로 나누어 제시하면 다음과 같다.

(단계 1) 작업하달 시점의 결정

작업하달을 위한 첫 단계로 작업하달 시점을 결정한다. 대사업체의 작업하달은 보통 8시간 마다의 교대작업 시점에 이루어진다는 점을 감안하여 작업하달은 8시간 마다의 이루어지는 FA 규칙을 사용하기로 한다. 작업하달 시점의 결정은 본 논문이 제시하고자 하는 주된 주제가 아니다.

(단계 2) 작업하달 로트의 결정

이 단계에서는 단계 1에서 결정된 시점에서 작업하달 할 대상 로트를 선정한다. 이 때, 고객의 오더가 웨이퍼를 담은 상자인 카세트 단위의 로트로 분할되어 작업하달 시점에 작업하달이 가능한 로트들에 대해서 작업하달 대기열에 여유시간이 적은 순으로 대기되어 있어야 한다.

(세부단계 1) 대기열에서의 대기시간과 여유시간의 평균과 분산의 수집

로트들이 기계 대기장소에서 떠날 때 각 로트가 경험하는 대기시간과 여유시간에 대한 평균과 분산의 각각의 추정값을 읽어 온다.

(세부단계 2) 작업장 로트들의 정보수집

작업하달 시점에서의 작업장의 모든 로트의 위치 및 상태에 대한 정보를 수집한다. 여기에는 각 기계 대기열별 대기중인 로트들과 그 로트들의 현재 가공공정의 번호, 각 기계별 가공중인 로트들에 대한 정보가 포함된다.

(세부단계 3) 로트들의 작업장에 가상투입

작업하달 시점에서 작업하달 대기열에서 대기중인 로트중 여유시간이 적은 순으로 $a \cdot \lambda$ (단, $a \geq 1$) 수만큼의 로트를 가상적으로 투입을 고려한다(가상으로 투입되는 로트를 가상로트

라 부르겠다). 참고로, a 의 값은 고정된 값으로 유지할 수도 있으나 동적으로 그 값을 결정하기 위해서는 현장 실무자가 작업장에 체류중인 로트수 등을 반영할 수 있다. 최적의 a 의 값을 결정하는 문제는 본 논문에서는 다루지 않기로 한다.

(세부단계 4) 가상로트들의 후속공정들에서의 시간대별, 기계별 작업부하의 예측

작업장에 체류중인 가상로트의 후속공정들에서의 도착시각을 예측한다. 본 논문에서 제시하는 후속공정에서의 도착시각을 계산하기 위하여 전진부하 방식을 채택한다. 그러나 본 논문에서는 기존의 연구들에서 채택하는 방식인 각 기계 대기장소에서의 평균대기시간을 사용하지 않고 각 기계 대기장소에서 작업장의 운영규칙인 최소여유시간 규칙을 반영하여 각각의 대기시간을 별도로 계산하여 로트가 후속기계에 도착하는 추정예상시각의 정확도를 높였다. 따라서 특정시점의 특정 기계에 걸리는 부하가 보다 정확하게 예측된다.

예를 들어, 어떤 로트가 기계 2, 기계 1, 기계2의 순서로 기계를 방문하여 가공된다고 가정하자. <그림 2(a)>에서 볼 수 있듯이 기존의 방식에서는 평균대기시간과 가공시간을 이용하여 후속기계에 도착하는 시각을 예측한다. 본 논문에서는 <그림 2(b)>에서 볼 수 있듯이 각각의 기계에서 로트들이 대기할 시간을 별도로 구하여 그 대기시간과 가공시간을 이용하여 후속 기계들에 도착하는 시각을 예측한다.

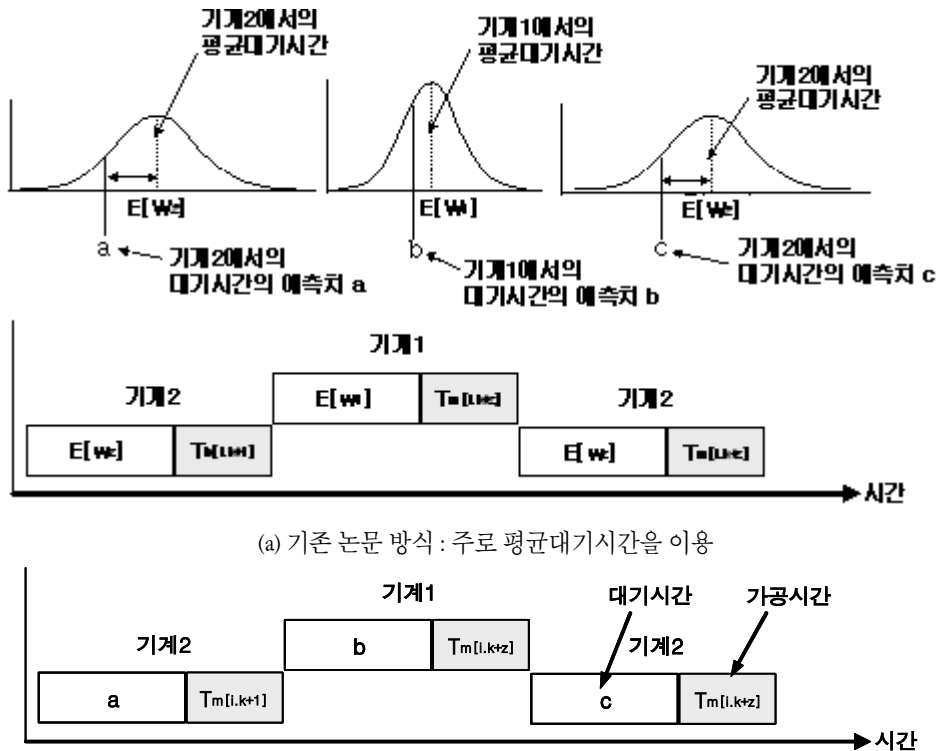
이상의 방식으로 작업하달 이후 시점에 대하여 작업장의 로트와 가상로트들이 각 기계에 도착하는 시점을 계산하여 시간대별 기계별로 작업부하를 누적한다. 이때, 작업장의 로트에 의한 예상부하와 가상로트에 의한 예상부하를 별도로 계산하여 저장해 둔다.

(세부단계 5) 작업부하에 의한 작업하달 판단

각 기계의 기준부하수준과 각 기계에서의 기계별, 시간구간별 예상부하를 비교하여 가상로트에 대한 승인 및 취소를 하는 단계이다. 예를 들어, 작업장의 로트와 가상로트에 의한 예상부하가 <그림 3(a)>의 기계 3($j=3$)과 같이 어느 한 기계 이상에서 계획기간의 모든 시간구간에 대해서 기준부하수준을 초과하면 가상로트의 일부분을 취소하여야 한다. 가상 투입된 로트를 취소하기 위해서는 세부단계 6로 간다. 그렇지 않으면 종료한다.

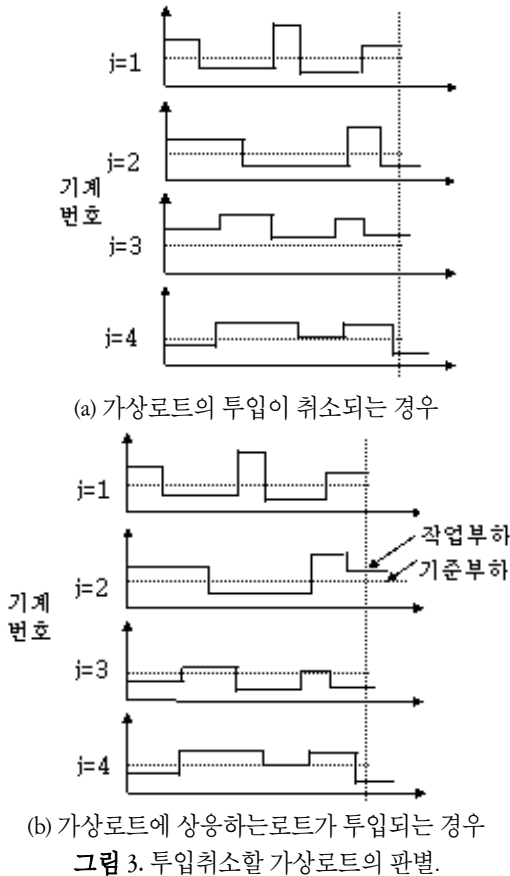
(세부단계 6) 투입을 취소할 가상로트의 선정

어느 한 기계 이상에서 계획기간 내의 모든 시간구간에 대해서 예상부하가 기준부하수준 초과하면 가상로트 중에서 투입을 취소할 로트를 선정한다. 이 절차의 기본생각은 계획기간의 모든 시간구간에 대해서 기준부하수준을 초과하는 예상부하를 가지는 기계를 방문하게 되는 가상로트들 중에 여유시간이 큰 순으로 투입을 취소하는 것이다. 예를 들어, 투입된 가



(a) 기존 논문 방식 : 주로 평균대기시간을 이용
 (b) 본 논문의 방식 : 별도로 대기시간을 계산
 그림 2. 로트의 각 기계의 도착시점의 추정하는 방식의 예.

상로트 중에서 여유시간이 큰 순으로 정렬한 것이 로트 1, 2, 3 이라 가정하자. 만약 가상으로 투입된 로트들에 의해서 <그림 3(a)>의 경우가 발생하면 로트1이 가장 먼저 투입이 취소할 가상로트로 선정된다. 가상로트 1이 취소될 로트로 선정된 후 만약 <그림 3(b)>의 경우가 발생하면 가상로트 2와 3은 투입을 취소할 로트로 선정되지 않는다. 따라서, 가상로트 1에 상응하는 실제 로트는 작업하달하지 않고, 가상로트 2와 3에 상응하는 실제 로트는 작업장에 투입한다.



(세부단계 8) 투입된 가상로트의 취소

투입을 취소할 가상로트로 선정된 것에 상응하는 작업하달 대기열의 실제 로트는 작업장에 투입하지 않는다. 세부단계 5로 간다.

4. 시뮬레이션 실험

작업하달 논리를 평가하기 위해 시뮬레이션 실험을 실시하였다. 시뮬레이션은 대상업체의 공정분석을 통해서 얻어진 시뮬레이션 모델과 데이터를 사용하였으므로 실제 반도체 제조업체의 현장을 잘 반영했다고 판단된다. 참고로, 본 연구의 시뮬레이션 실험에 사용된 모델과 입력 데이터는 선행연구인 김기영 외 3인(2005)의 논문내용을 개선시킨 것이므로 주문형 반도체

웨이퍼 제조업체의 공정분석과 시뮬레이션에 관심이 있는 독자는 이 논문을 참조하면 된다.

본 논문에서 제안한 작업하달 방식(이후 편의상 OSAR이라 부르겠다)을 평가하기 위해서 Kim et.al.(1996)의 SA based projected Queue방법(이후 편의상 SA2라 부르겠다)과 비교하였다. 시뮬레이션 실험은 시뮬레이션 전용 패키지인 eM-Plant5.5를 이용하여 518400분 동안 수행되었으며, 실험별 반복횟수는 10회로 하였다.

<그림 4>는 지연된 로트의 수의 비율을 비교한 그래프인데, 로트 유형에 무관하게 SA2보다 OSAR 경우에서 더 낮은 지연된 로트의 수의 비율을 보여주고 있다. 특히, 작업장의 모든 기계에서 거치는 로트 유형 PT1에 대해서는 SA2 보다 OSAR 경우에서 비교적 낮은 지연된 로트의 수의 비율을 보이고 있다.

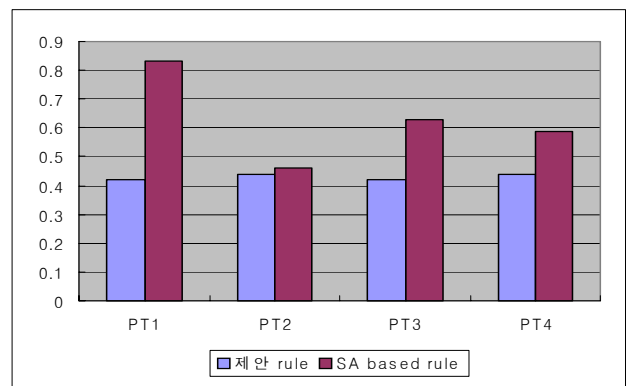


그림 4. 지연된 로트 수의 비율의 비교.

<그림 5>은 시간이 경과됨에 따른 생산된 로트 수를 비교한 것이다. 시간이 경과될수록 SA2 보다 OSAR 경우에서 생산된 로트가 더 커다는 것을 보여주고 있다. 반도체 생산업체에서 산출율이 중요하는 것을 감안한다면 OSAR을 사용하는 것이 더 유리할 수 있음을 알 수 있다.

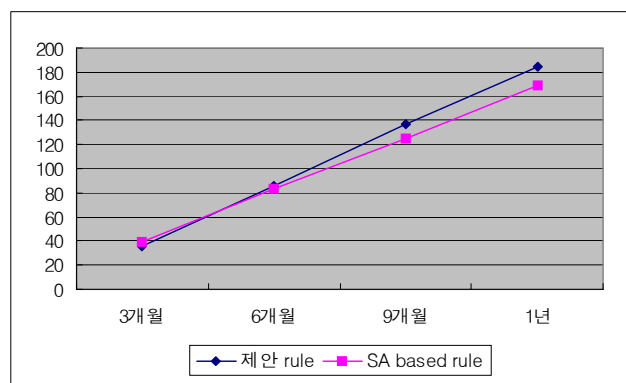


그림 5. 시간의 경과에 따른 생산된 로트 수의 비교.

<그림 6>는 주요 기계들에서 이용율을 도시한 것이다. 여기서, SA2 보다 OSAR 경우에서 기계 이용율을 더 높음을 알 수

있다. SA2는 작업할 로트가 없어서 기계를 가동하지 못하는 상황을 피하는 것이 기본생각이고 OASR은 작업할 로트에 의해 과도하게 부하가 걸리는 기계를 피하는 것이 기본생각이기 때문에 나온 결과라고도 볼 수 있다.

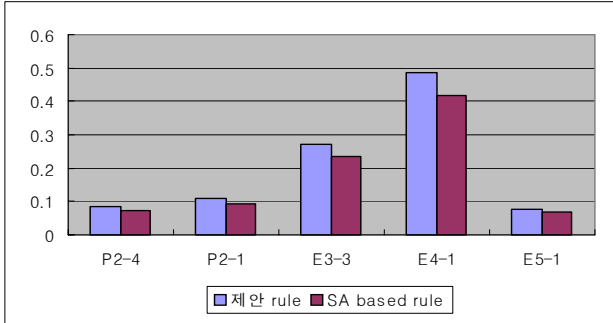


그림 6. 주요 기계에서의 이용률의 비교.

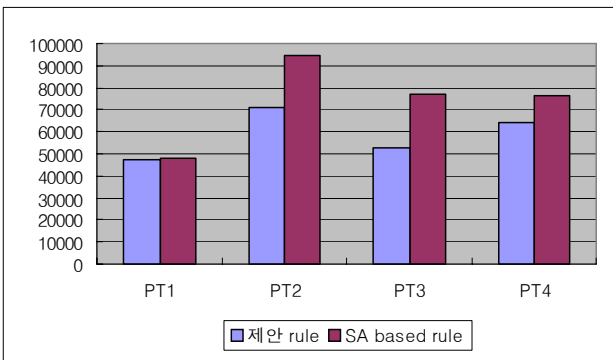


그림 7. 평균체류시간의 비교(단위: 분).

<그림 7>은 로트들의 평균체류시간을 비교한 그래프이다. SA2 보다 OASR 경우에서 평균체류시간이 더 적음을 보여주고 있다. 로트 유형 PT1에 대해서는 OASR과 SA2는 평균체류시간이 비슷한 반면에 로트 유형 PT2, 3, 4에 대해서는 SA2 보다 OASR 경우에서 평균체류시간의 더 적다. PT1은 모든 기계를 거치는 반면 PT2, PT3, 그리고 PT4는 가공을 위해서 이미 거쳤던 기계에 다시 방문하는 경우가 많기 때문이라 생각된다. 따라서 재가공정이 많은 로트들이 가공된다면 SA2 보다 OASR을 사용하는 경우가 더 유리할 수 있다.

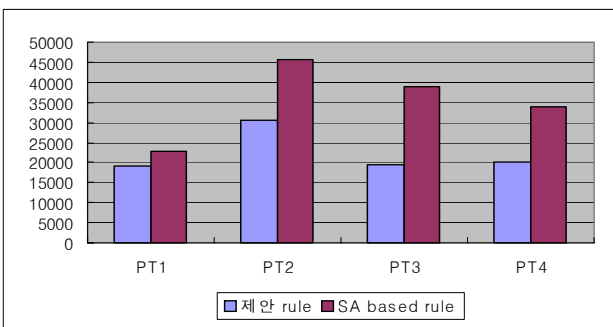


그림 8. 평균 lateness의 절대치(단위: 분).

<그림 8>은 평균 lateness의 절대치를 비교한 그래프인데, SA2보다 OASR 경우에서 평균 lateness의 절대치들이 적음을 보여주고 있다. 즉, SA2보다 OASR을 사용하는 경우가 더 납기일 맞도록 로트들을 생산할 수 있음을 알 수 있다.

5. 결론

본 연구에서는 주문형 반도체를 생산하는 회사를 대상으로 디스패칭규칙이 최소여유시간 규칙일 때, 작업장의 공정재고를 최적으로 유지하기 위한 작업하달 논리를 개발하였다.

개발된 작업하달 논리는 다음과 같은 특징을 가지고 있다. 첫째, 평균적으로 투입되는 작업하달 로트의 수보다 조금 더 많은 로트를 가상적으로 투입하여 각 시점별, 기계별로 작업 부하를 계산한 후, 부하가 과다하면 투입된 로트를 취소해 가는 방식이다. 둘째, 각 시점별, 기계별로 작업부하가 걸리는 시점을 보다 정확하게 계산하기 위해서 각 기계의 대기장소에서 우선순위규칙인 로트의 최소여유시간 규칙을 반영하여 로트별 시점별, 기계별로 대기시간을 별도로 예측하는 방식을 사용하였다.

시뮬레이션 실험은 본 연구에서 개발된 작업하달 논리가 기존의 작업하달 논리보다 다양한 평가측도에서 좋은 성능을 보임을 알 수 있다. 따라서 주문형 반도체 제조회사의 작업하달 논리로 사용한다면 좋은 성능을 발휘할 수 있다고 생각된다.

감사의 글

본 연구는 국책국가연구개발사업/엔지니어링핵심기술 사업의 APS로직 설계 및 시뮬레이션 개발이란 과제의 연구비로 수행되었습니다. 그리고 본 논문의 개선을 위해서 자세하게 심사해주시고 많은 조언을 주신 두 심사자님께 진심으로 깊은 감사를 드립니다.

참고문헌

김기영, 이정호, 강창호, 김갑환, 주문형 반도체 웨이퍼 공정분석을 위한 시뮬레이션 연구, *IE Interface*, 18(1), 22-34, 2005.
 Corsten, H., R. Gossinger and N. Wolf, Flexibility-driven Order Releases in Job-Shop Production, *Technovation*, 25, 815-830, 2005.
 Dabbas, R. M., H. N. Chen, J. W. Fowler, and D. Shunk, A Combined Dispatching Criteria Approach to Scheduling Semiconductor Manufacturing System, *Computers & Industrial Engineering*, 39, 307-324, 2001.
 Donselaar, K. H. and B. J. Gubbels, How to Release Orders in order to Minimize System Inventory and System Nervousness?, *International Journal of Production Economics*, 78, 335-343, 2002.
 Huang, M., D. Wang and W. H. Ip, Simulation Study of CONWIP for a Cold Rolling Plant, *Int. J. Production Economics*, 54, 257-266, 1998.

Kang, J., A Method for Target Scheduling of Semiconductor Wafer Fabrication Based on Event-Based Optimization Modeling and Discrete Event Simulation, Ph.D. thesis, University of California, Berkeley, 1996.

Kim, J., R. C. Leachman, and B. Suh, Dynamic Release Control Policy for the Semiconductor Wafer-fabrication Lines, *JORS*, 47, 1516-1525, 1996.

Kim, K. Y., K. H. Kim, C. H. Kang, Estimating Lead Times of Orders in a Semiconductor Wafer Manufacturer, The Proceedings of the 33th International Conference on Computers and Industrial Engineering, CD CIE024, 6 pages, Jeju, Korea, March 25-27, 2004.

Kingsman, B. G., Modeling Input-Output Workload Control for Dynamic Capacity Planning in Production Planning Systems, *International Journal of Production Economics*, 68, 73-93, 2000.

Koh, S.C.L. and S.M. Saad, The Use of Intelligent Feedback for Work Order Release in an Uncertain Manufacturing System, *Robotics and Computer-Integrated manufacturing*, 20, 517-527, 2004.

Ooijen, H.P.G. and J.W.M. Bertrand, Economic Due-date Setting in Job-Shop based on Routing and Workload Dependent Flow Time Distribution Function, *International Journal of Production Economics*, 74, 261-268, 2001.

Oosterman, B., M. Land, and G. Gaalman, The Influence of Shop Characteristics on Workload Control, *International Journal of Production Economics*, 68, 107-119, 2000.

Land, M. and G. Gaalman, Workload Control Concepts in Job Shops: A Critical Assessment, *International Journal Production Economics*, 46-47, 535-548, 1996.

Lee, Y. S., S. Y. Kim, S. H. Yea, and B. K. Kim, "Production Planning in Semiconductor Wafer Fab Considering Variable Cycle Times," *Computers ind. Engng*, 33(3-4), 713-716, 1997.

Lin, C. Y., Shop Floor Scheduling of Semiconductor Wafer Fabrication Using Real-Time Feedback Control and Predictions, doctoral dissertation, University of California at Berkeley, 1996.

Raaymakers, W. H. M., J. W. M. Bertrand, and J. C. Franscoo, Makespan Estimation in Batch Process Industries Using Aggregate Resource and Job Set Characteristics, *International Journal of Production Economics*, 70, 145-161, 2001.

Shen, Y., Stochastic Wafer Fabrication Scheduling, Ph.D. thesis, University of California, Berkeley, 1997.

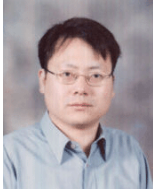
Sivakumar, A. I. and C. S. Chong, A Simulation based Analysis of Cycle Time Distribution, and Throughput in Semiconductor Backend Manufacturing, *Computers in Industry*, 45, 59-78, 2001.

Sloan, T. W., Scheduling Semiconductor Wafer Manufacturing Using In-Line Equipment Condition and Yield Information, Ph.D. thesis, University of California, Berkeley, 1998.

Sung, C. S. and Y. I. Choung, Minimizing Makespan on a Single Burn-in Oven in Semiconductor Manufacturing, *European Journal of Operational Research*, 120, 559-574, 2000.

Vargas-Villamil, F. D. and D. E. Rivera, Multilayer Optimization and Scheduling Using Model Predictive Control: Application to Reentrant Semiconductor Manufacturing Lines, *Computers and Chemical Engineering*, 24, 2009-2021, 2000.

Weng, Z. K., Manufacturing Lead Times, System Utilization Rates and Lead-Time Related Demand, *European Journal of Operational Research*, 89, 259-268, 1996.



김기영

부산대학교 산업공학과에서 공학사, 석사, 박사
 현재: 동서대학교 부교수
 관심분야: 물류시스템



강창호

부산대학교 산업공학 석사(2005년)
 동부정보기술(주) 물류운영팀 소속
 2005년~현재: 동부부산컨테이너터미널
 운영 업무 수행



김갑환

부산대학교 산업공학과 교수로 근무중
 한국과학기술원에서 석사, 박사학위를 취득
 생산관리, 물류시스템 분야를 강의하고 연구