行政院國家科學委員會專題研究計畫　成果報告

半導體廠小批量報廢決策(I)

計畫類別：個別型計畫
計畫編號：NSC91-2213-E-009-134-
執行期間：91 年 08 月 01 日至 92 年 07 月 31 日
執行單位：國立交通大學工業工程與管理學系

計畫主持人：巫木誠

報告類型：精簡報告

處理方式：本計畫可公開查詢

中　華　民　國 92 年 8 月 25 日

摘要—半導體廠小量晶圓批是否應該繼續加工或逕行報廢，是經常面臨的一個決策問題，這個問題過去的學術文獻相關的研究相當少，工業界對此通常是以直覺方式處理，缺乏有系統的分析方法。本研究構建一決策模式並提出一有系統的演算法，以決定小量晶圓批在各加工道次的報廢法則，使晶圓廠的利潤最大化。根據模擬實驗結果，在低良率高單價的情境下，本研究所提的報廢方法，可提高晶圓廠的利潤高達23%。
關鍵詞：小量晶圓批、瓶頸、低良率、高單價、新產品導入

Abstract—Some wafers in a lot may become spoiled after they are processed at a workstation; such a lot is called a small lot. In a low yield and high price scenario, scrapping small lots may increase revenue and profit; yet this notion has seldom been examined. This study presents a model for formulating the decision problem of scrapping small lots. A genetic algorithm is used to solve the problem when the solution space is large. An exhaustive search method is used when the solution space is small. Some numerical examples are used to evaluate the outcome of scrapping small lots. The profit obtained by the proposed scrapping method may be up to 23% higher than that obtained without scrapping.

Keywords: small lots, bottleneck, low yield, high price, product introduction

I. INTRODUCTION

A wafer in a semiconductor fab is transported in a fixed-size batch. Such a batch is called a *wafer lot* (or a *lot*) that normally includes 25 wafers. Due to yield problem, some wafers in a lot during processing may become spoiled and cannot be processed further. The number of good wafers in a lot, less than 25, is called a *small lot*, while that includes 25 good wafers is called a *full lot*.

The manufacturing cost per wafer for a small lot, in batch-typee workstations, is higher than that of a full lot. Machines in a fab are generally classified into two types, *series-type* and *batch-type*. A batch-type machine in one run simultaneously processes a batch up to six lots. The running cost per wafer for a small lot on a batch-type machine is therefore higher than that for a full lot. A series-type machine in one run processes a single wafer and its running cost per wafer is independent of lot size.

A semiconductor fab may face a decision problem about the scrapping of small lots. For example, given a small lot of 12 good wafers and with ten layers remaining to be processed, should the fab keep the lot for further processing or scrap it? Keeping the small lot until its completion will create revenue, while scrapping the lot provides an opportunity for processing new full lots. These cost/revenue factors should all be included when making the decision to scrap.

Much literature on semiconductor yield modeling and its applications has been published [1]. Yet, very few study the decision for scrapping small lots in semiconductor manufacturing. Daigle and Powell propose a formalized management procedure to reduce wafer scraps [2]. Based on the cost of yield, Maynard et al. proposed a heuristic method in IBM for the scrap decision of wafers [3-4]. Interviews from industrial workers reveal that the decision of scrapping small lots is often made heuristically.

This paper develops a mathematical model for the decision problem of scrapping small lots in a semiconductor fab. Based on the model, a genetic algorithm is proposed for making the scrapping decisions at each layer. However, when the number of low-yield layers is few, the exhaustive search method is used to determine the associated scrapping rules. Simulation experiments show that scrapping small lots as proposed may considerably increase profit.

II. MODEL

The semiconductor fab of interest produces only one product and involves two types of workstation, the series-type and the batch type. Let BT_s represent the bottleneck of the series-type workstations and BT_b represent the bottleneck of the batch-type workstations.

Parameters

L: total number of layers

M: total number of wafers in a full lot

AT_s: capacity (available run time) of BT_s

AT_b: capacity (available run time) of BT_b

ts_i: required run time of an operation processed by BT_s at layer i; $0 \le i \le L$

tb_i: required run time of an operation processed by BT_b at layer i; $0 \le i \le L$

n: number of lots simultaneously processed by BT_b

P: price of the product

FC: fixed cost of the fab

$\overline{C_i} = \left[c_k^i\right]^T$: processing cost per lot at layer i, $0 \le k \le M$; $0 \le i \le L$

c_k^i : processing cost for a lot with k wafers at layer i.

$A_i = \left[a_{jk}^i\right]$: yield matrix at layer i, $0 \le j \le M$; $0 \le k \le M$; $0 \le i \le L$

a_{jk}^i : probability that a lot with j wafers becomes one with k wafers, after completing the operations at layer i.

if $j \ge k$ then $1 \ge a_{jk}^i \ge 0$

if $j < k$, then $a_{jk}^i = 0$

$$\sum_{k=0}^{M} a_{jk}^{i} = 1$$

Variables

$\overline{U_i} = \left[u_k^i\right]$: distribution of output lots at layer i when only one lot is released to the fab, $0 \le k \le M$, $0 \le i \le L$; $\overline{U_0} = [1, 0, ..., 0]$

u_k^i: number of output lots that carry k wafers at layer i when only one lot is released to the fab

$\overline{W_i} = \left[w_k^i\right]$: distribution of output lots at layer i, $0 \le k \le M$, $0 \le i \le L$

w_k^i: number of output lots that carry k wafers at layer i

$S(\overline{W_i}) = \sum_{k=1}^{M} k \cdot w_k^i$: total number of output wafers at layer i

$L(\overline{W_i}) = \sum_{k=1}^{M} w_k^i$: total number of output lots at layer i

$\overline{h} = [h_i]$: decision vector for scrapping small lots

h_i: threshold for scrapping small lots at layer i, $1 \le i \le L$ and $1 \le h_i \le M - 1$.

$R(h_i) = \left[r_{jk}^i\right]$: scrapping matrix at layer i, $0 \le j \le M$; $0 \le k \le M$; $1 \le i \le L$

r_{jk}^i: a binary number (0 or 1) which indicates the probability that a lot with j wafers becomes one with k wafers at layer i, when the scrapping rule (h_i) is implemented

If $j > h_i$ and $k = j$ then $r_{jk}^i = 1$;

If $j > h_i$ and $k \ne j$ then $r_{jk}^i = 0$;

If $j \le h_i$ and $k = 0$ then $r_{jk}^i = 1$;

If $j \le h_i$ and $k \ne 0$ then $r_{jk}^i = 0$;

$$\sum_{k=0}^{M} r_{jk}^{i} = 1$$

$a(\overline{h})$: number of input lots that can fully utilize the bottleneck of the fab, when $\overline{h}$ is applied in the fab.

The decision problem of scrapping small lots can be mathematically formulated as follows.

Max $\quad P \cdot S(\overline{W_L}) - \left(\overline{W_0} \times \overline{C_0} + \sum_{i=1}^{L} \overline{W_{i-1}} \times \overline{C_i}\right) - FC$

Subject to

$$\overline{U_i} = \overline{U_0} \times \prod_{1}^{i} A_i \times \prod_{1}^{i} R(h_i) \qquad (1)$$

$$S(\overline{U_i}) = \sum_{k=1}^{M} k \cdot u_k^i \qquad (2)$$

$$L(\overline{U_i}) = \sum_{k=1}^{M} u_k^i \qquad (3)$$

$$a(\bar{h}) = Min\{\frac{AT_s}{\sum_{i=0}^{L-1} ts_i \cdot S(\overline{U_i})}, \frac{AT_b}{\sum_{i=0}^{L-1} tb_i \cdot \frac{L(\overline{U_i})}{n}}\} \qquad (4)$$

$$\overline{W_i} = a(\bar{h}) \cdot \overline{U_i} \qquad (5)$$

$$S(\overline{W_L}) = \sum_{k=1}^{M} k \cdot w_k^L \qquad (6)$$

$$h_i \geq h_j \quad \text{if } i < j \qquad (7)$$

The objective function is to maximize the profit. In Eq. (1)-(3), $\overline{U_0}$ refers to the input in the scenario where only one lot is released to the fab; $\overline{U_i}$ represents the consequent output at layer *i* after the yield problem and scrapping rules are addressed. Eq. (4) gives the maximum number of lots $a(\bar{h})$ that must be released to the fab to utilize the capacity fully. Eq. (5) determines $\overline{W_i}$, the output at layer *i* when $a(\bar{h})$ lots are released to the fab. Eq. (6) determines $S(\overline{W_L})$ the total number of wafers output by the fab.

Eq. (7) denotes that the scrapping threshold at an upstream layer should not be smaller than that at a downstream layer. Otherwise, it is an irrational decision. For example, a case with $h_1 = 2$ and $h_2 = 4$ means that a small lot with 3 wafers will pass layer 1 but will be scrapped at layer 2. This implies that the processing of this lot at layer 2 is useless. That is, even the yield of this lot at layer 2 is 100%, the lot should still be scrapped.

The objective function is a quite complex nonlinear function. Observing these complex properties, we proposed the use of a genetic algorithm to solve the problem when the solution space is large. However, when the solution space is small (for example, including three critical layers or less), an exhaustive search is performed to solve the problem.

III. EXPERIMENT RESULTS

Some experiments using either GA or exhaustive search have been carried out. The fab of interest produces only one product, with 20 layers in process route. Of the 20 layers, those with 100% yield are called *non-critical layers*; the others are called *critical layers*. The binomial distribution is used to estimate the yield matrix associated with a critical layer. Experiments for various numbers of critical layers are tested.

Table 1 summarized the experimental results, with only one critical layer. For each *CL*, the fab bottleneck, the number of input lots, the revenue, the variable costs, and the profit are presented. The profit would improve up to 23.8%, comparing to the case without scrapping small lots.

IV. CONCLUSION

This study formulates a model for solving the decision-making problem concerning the scrapping of small lots in semiconductor wafer fabs. When the number of low-yield layers is less than or equals three, the exhaustive search method is suggested to solve the formulated problem. Otherwise, GA (genetic algorithm) is suggested. The proposed scrapping method considerably outperforms the no-scrapping method when the critical layers are in the upstream. Yet, there may be no difference when the critical layers are in the downstream.

Solutions of the numerical examples reveal the following two interesting phenomena concerning a low yield fab. First, given a single process route, the bottleneck of the fab may switch between a series-type workstation and a batch-type workstation. Second, the difference between the profit obtained by scrapping and that obtained without scrapping is substantial when the low yield layers are upstream.

References

1. W. Kuo and T. Kim, “An Overview of Manufacturing Yield and Reliabilty Modeling for Semiconductor Products,” *Proceedings of the IEEE*, pp. 1329-1344, 1999.
2. K. Daigle and R. Powell, “Manufacturing Scrap Reduction Team,” *IEEE/SEMI Advanced Manufacturing Conference*, pp. 230-231, 1996.
3. D. N. Maynard, D. S. Kerr, and C. Whiteside, “Determining cost of yield to monitor fab manufacturing processes,” *micromagazine.com*, pp. 63-69, June 2003.
4. D. N. Maynard, D. S. Kerr, and C. Whiteside, “Cost of Yield,” *IEEE/SEMI Advanced Manufacturing Conference*, pp. 165-170, 2003.

Table 1: Apply optimum scrap rules to scenario S1

CL	$a_s(\bar{h})$ (lots)	$a_b(\bar{h})$ (lots)	$a(\bar{h})$ (lots)	h_{CL+1}	Output wafer	TR ($M)	TVC ($M)	FC ($M)	Profit ($M)	Case 1 Profit ($M)	Profit improvement (%)
1	5894	3520	3520	9	23650	68.6	22.9	24.0	21.7	17.5	23.8%

2	4605	2786	2786	8	22504	65.2	21.4	24.0	19.8	17.3	14.6%
3	3936	2415	2415	7	21833	63.3	20.8	24.0	18.5	17.0	8.8%
4	3676	2392	2394	7	21643	62.7	21.0	24.0	17.8	16.8	5.8%
5	3346	2223	2223	6	21332	61.8	20.7	24.0	17.1	16.6	3.4%
6	3165	2214	2214	6	21255	61.6	20.9	24.0	16.7	16.3	2.2%
7	2968	2141	2141	5	21112	61.2	20.9	24.0	16.3	16.1	1.1%

計畫成果自評

(1) 本研究成果已完全達成原計畫書所規劃之目標

(2) 本研究部分成果已經投稿至 *IEEE Transactions on Semiconductor Manufacturing*，3 位 reviewer 均對原創性表示肯定，但對細節有些建議，目前已針對 reviewer 的意見修改，重新送審。

(3) 本研究部分成果已經發表於 M. C. Wu, C. W. Chiou, and H. M. Hsu, “Scrap Rules For Small Lots in Wafer Fabrication,” 2002 Semiconductor Manufacturing Technology Workshop, Hsin-Chu, Taiwan, pp. 181-184, 2002.

可供推廣之研發成果資料表

□ 可申請專利 ☑ 可技術移轉 日期：91 年 8 月 20 日

國科會補助計畫	計畫名稱：半導體小批量報廢決策 計畫主持人：巫木誠 計畫編號：NSC91-2213-E-009-134 學門領域：工業工程與管理
技術/創作名稱	半導體小批量報廢決策演算法
發明人/創作人	巫木誠、邱志文
技術說明	中文：半導體廠小量晶圓批是否應該繼續加工或逕行報廢，是經常面臨的一個決策問題，這個問題過去的學術文獻相關的研究相當少，工業界對此通常是以直覺方式處理，缺乏有系統的分析方法。本研究構建一決策模式並提出一有系統的演算法，以決定小量晶圓批在各加工道次的報廢法則，使晶圓廠的利潤最大化。根據模擬實驗結果，在低良率高單價的情境下，本研究所提的報廢方法，可提高晶圓廠的利潤高達 23%。
	英文：Some wafers in a lot may become spoiled after they are processed at a workstation; such a lot is called a small lot. In a low yield and high price scenario, scrapping small lots may increase revenue and profit; yet this notion has seldom been examined. This study presents a model for formulating the decision problem of scrapping small lots. Some numerical examples are used to evaluate the outcome of scrapping small lots. The profit obtained by the proposed scrapping method may be up to 23% higher than that obtained without scrapping.
可利用之產業及可開發之產品	半導體晶圓廠
技術特點	低良率、高單價情境提高利潤的決策方法
推廣及運用的價值	在低良率高單價情境下可提高利潤高達 24%

1. 每項研發成果請填寫一式二份，一份隨成果報告送繳本會，一份送貴單位研發成果推廣單位（如技術移轉中心）。

2.本項研發成果若尚未申請專利，請勿揭露可申請專利之主要內容。

3.本表若不敷使用，請自行影印使用。