

Semiconductor Product-mix Estimate with Dynamic Weighting Scheme

Argon Chen, Ziv Hsia and Kyle Yang
 Graduate Institute of Industrial Engineering, National Taiwan University
 1 Roosevelt Rd. Sec. 4, Taipei, Taiwan, 106
 achen@ntu.edu.tw

Abstract – Semiconductor manufacturing technology is progressing rapidly. As shown by the celebrated Moor's laws, new generations of semiconductor products emerge in the pace of every 6 months. As a result, technologies of different generations often co-exist in the same fabrication facility. For effective manufacturing resource planning, accurate prediction of product mix is therefore crucial. Product-mix prediction is usually made by performing the demand disaggregation function in demand planning. In this paper, weighted product-mix estimation methodologies are first proposed. Dynamic weighting schemes are then developed to improve the accuracy of product mix prediction. The methodologies will be tested with simulated DRAM demands and actual semiconductor demands of different technology generations.

INTRODUCTION

Results of demand planning serve as the basis of every planning activity in the supply network and ultimately determine the effectiveness of manufacturing/logistic operations in the network. A common demand planning approach is to make forecast at an aggregated level first, and then break down (disaggregate) the forecast statistically and/or judgmentally into the individual forecasts. This is known to be an effective means for better forecasting because the aggregated product-group demand is observed to fluctuate less and easier to make forecast. This approach, however, requires an accurate product-mix prediction to break down the aggregated demand forecast into individual product forecasts. Product-mix can be expressed as demand proportions in a product group. For example, proportions of 64M, 128M and 256M DRAM demands form the product-mix of the DRAM product group. Thus, product-mix estimation is equivalent to demand disaggregation. That is, to estimate the product-mix, we need to estimate the proportions of individual demands in the product group.

An effective product mix estimate should take into consideration the characteristics of semiconductor demands. There are three important characteristics that will be considered in our proposed methodologies:

1. Product substitutability: most products within the same product group are substitutable among one another.
2. Product life cycle (PLC): Fast-paced technology development leads to fast PLC transition.
3. Variability proportional to volume: the greater the demand volume, the more volatile the demand.

In Fig. 1, we simulate the 128Mb, 256Mb and 512Mb DRAM demands to resemble these three characteristics.

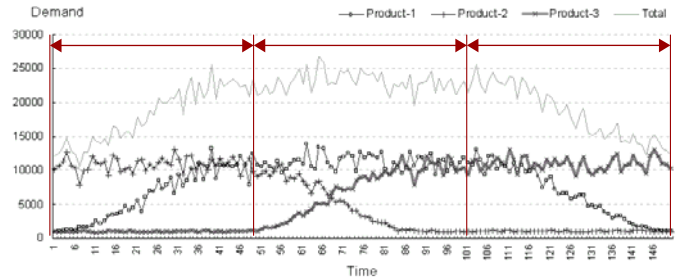


Fig. 1 Simulated DRAM demands

It can be seen, substitution of phasing out product demand by the emerging product leads to dramatic change of product-mix. It is difficult to estimate the product-mix without considering the product life cycles and substitutability.

In this paper, we first propose weighted product-mix estimation methodologies. Dynamic weighting schemes are then developed to more effectively capture the PLC for better product-mix prediction. Finally, both simulated DRAM demands and actual semiconductor demands of different technology generations are used to test the proposed methodologies.

WEGHITED PRODUCT-MIX ESTIMATION

Among the many product-mix estimates (or demand disaggregation methods) [1], the following two methods are most widely used:

$$\text{Method A: } \hat{p}_i = \frac{\sum_{t=1}^n d_{it}}{\sum_{t=1}^n D_t}$$

$$\text{Method B: } \hat{p}_i = \frac{\sum_{t=1}^n d_{it}}{n} \bigg/ \frac{\sum_{t=1}^n D_t}{n} \quad (1)$$

where d_{it} is the demand of product i at period t ; n is the number of historical demand periods; $D_t = \sum_{i=1}^k d_{it}$ is the total demand at period t ; k is the number of products in the product group; and \hat{P}_i is the mean-proportional estimate of product i .

As mentioned earlier, directly forecasting on individual product demands usually result in a far-off forecast that not only impairs the quality of subsequent manufacturing plans but also send the ripples to the down-stream manufacturing activities via the supply chain. To capture the effect of PLC change on the product-mix, we would give higher weights to the more recent data; that is, we like to have the product-mix estimate more influenced by more recently observed demands. Thus, we apply exponential weights to the product-mix estimates in (1), (7), and (13), respectively, to obtain product-mix prediction for the next time period. First, the mean-proportional estimate in (1) now becomes an exponentially weighted average of historical demands:

$$\hat{p}_{i,n+1} = \frac{\sum_{t=1}^n w_{it} \cdot d_{it}}{\sum_{j=1}^m \sum_{t=1}^n w_{jt} \cdot d_{jt}} \quad (2)$$

where $\hat{p}_{i,n+1}$ is the prediction of product i proportion for period $n+1$; and w_{it} is the weight applied to product i demand at time t and satisfies:

$$\sum_{t=1}^n w_{it} = \sum_{t=1}^n \frac{\alpha_i (1 - \alpha_i)^{n-t}}{1 - (1 - \alpha_i)^n} = 1 \quad (3)$$

with the constant α_i to control the declining rate over time as shown in Fig. 2.

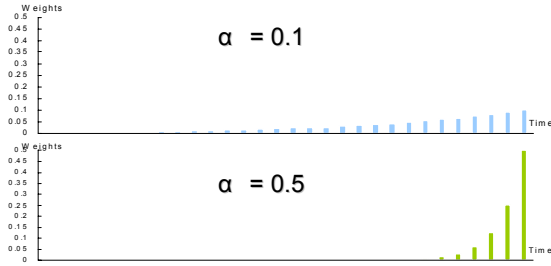


Fig. 2 Exponential weights controlled by α values

How to choose appropriate α_i values in (3) becomes critical for accurate product-mix estimate.

PLC LEADING INDICATOR

To find the best smoothing constants α_i of forecast estimate at period c , that is to find smoothing constants α_i ($i=1 \sim k$), we minimize the mean squared forecast errors over s periods:

$$SSE(s) = \sum_{t=k+1}^{k+s} \sum_{i=1}^k (\hat{p}_{i,t} - p_{i,t})^2 \quad (4)$$

Determination of good α_i is a very time consuming task. Suppose each α_i has 99 possible values ($\alpha_i = 0.01 \sim 0.99$). Then, there are 99^k possible $\{\alpha_1, \alpha_2, \dots, \alpha_k\}$ candidates. The

time to compute all candidates to search for the best $\{\alpha_1, \alpha_2, \dots, \alpha_k\}$ to minimize (4) requires enormous computing power. The most common method is Steepest Descent Search [2]. The search of the smoothing constants can be treated as a k -variable ($\alpha_1 \sim \alpha_k$) steepest-descent search with (4) as the objective function to minimize. Treat $\{\alpha_1, \alpha_2, \dots, \alpha_k\}$ as a point " α " in k -dimension coordinate system. The steepest-descent search process here can be seen as a process to find the point in the k -dimension coordinate system. Furthermore, the movements of α from one point to another are reached by adjusting each α_i with fixed step size " r ". There are three possible movements for each α_i , adding one step size to α_i , ($\alpha_i + r$), subtracting one step size from α_i , ($\alpha_i - r$), or keep the α_i at the original position. The successive points of α representing the most SSE-reduced direction which is called the "gradient vector" while all the other possible directions are called "candidate vectors". Termination of the search occurs at the point where the gradient vector becomes null; i.e., the current α point has the smallest SSE value. An example of two-product steepest-descent smoothing constant search process is shown in Fig. 3.

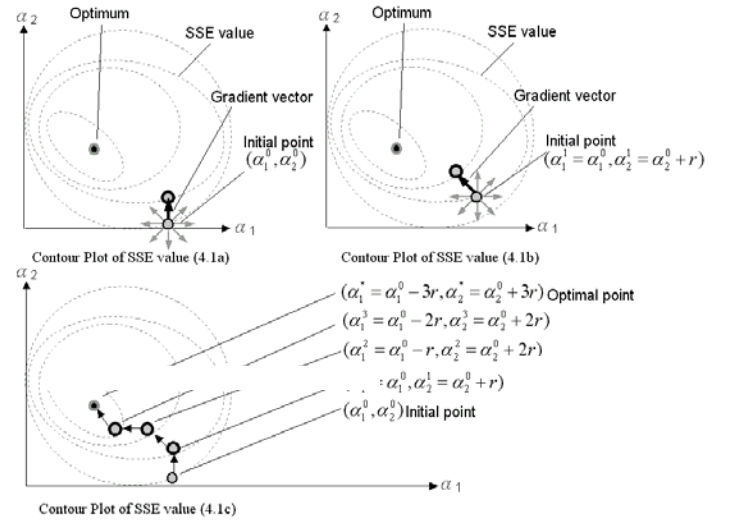


Fig.3 Two-product steepest-descent search process

The steepest descent method is still a computation-intensive method. It would be much more efficient if we know the direction toward the minimum. Taking the effect of changing demand variability into account, we invented a PLC transition leading indicator using the sample one-lag autocorrelation (SAC) statistic:

$$SAC_t = \hat{\rho} = \frac{\frac{1}{s} \sum_{\tau=t-1}^{t-(s-1)} (d_\tau - \bar{d})(d_{\tau+1} - \bar{d})}{\frac{1}{s} \sum_{\tau=t}^{t-(s-1)} (d_\tau - \bar{d})^2} \quad (5)$$

where SAC_t is the one-lag sample autocorrelation calculated at period t using demand dataset $\{d_{t-s+1}, \dots, d_t\}$. When the product is at the "growth" or "decline" phase, the product-mix proportion significantly rises or falls and the SAC becomes higher as well. If the product is mature in the market and its

proportion is stable, SAC will be lower because noise dictates the changes of the product-mix proportion (Fig. 4).

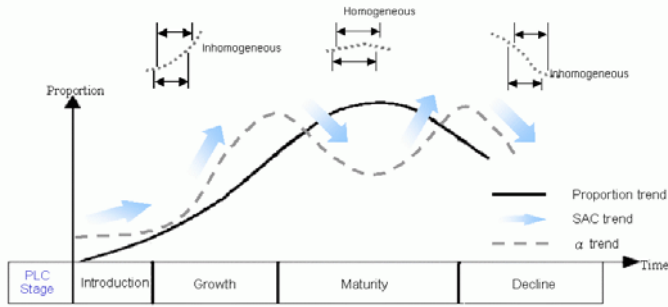


Fig. 4 Relationship among α , SAC and PLC

Sample size s in (5) determines how sensitive the SAC is to the PLC transition and to the demand noise. Fig. 5 shows the SAC values with sample sizes $s=15, 25$ and 50 for a simulated DRAM demand.

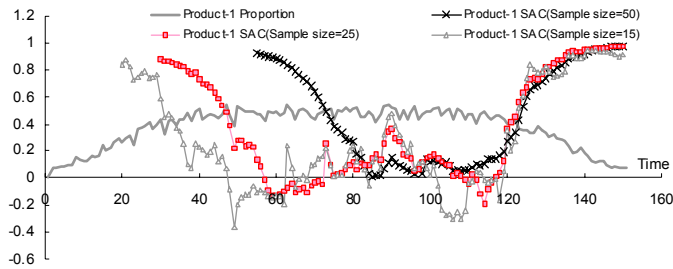


Fig. 5 SAC calculated by different sample sizes

It can be seen that the SAC with a large sample size ($s=50$) is too slow to reflect the PLC transition while the SAC with small sample size ($s=15$), though responsive to the PLC transition, is too sensitive to the demand noise. The sample size of 25, approximately on half of one PLC phase, gives the SAC a very good indication of PLC transition.

Fig. 6 shows the smoothing constant estimated by steepest descent search (SDS), SAC calculated with $s=25$ and the product-mix proportions of the simulated DRAM data (Fig. 1). The smoothing constant estimate goes up when the trend is rising or declining but goes down in the maturity phase. The trend of SAC and the trend of SDS-estimated smoothing constants match each other pretty well. Therefore, SAC would be a good leading indicator to determine the changing trend of the smoothing constants. That is, the smoothing constant estimated at period $t+1$ should be higher than that at period t when SAC_{t+1} is higher than SAC_t and vice versa. Using the SAC as a leading indicator for the search for the best smoothing constants has been proven to cut the computing time to 1/55. Fig. 7 shows the procedure of the prediction scheme.

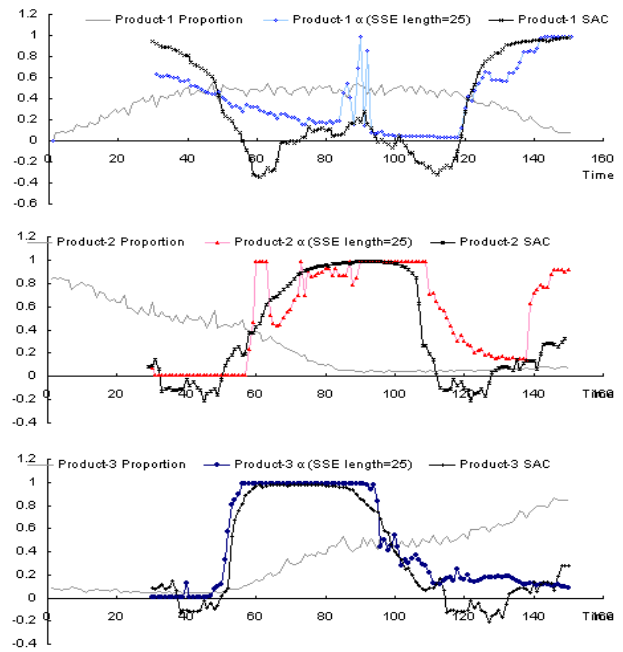


Fig. 6 Smoothing constant estimates and SAC_t (DRAM)

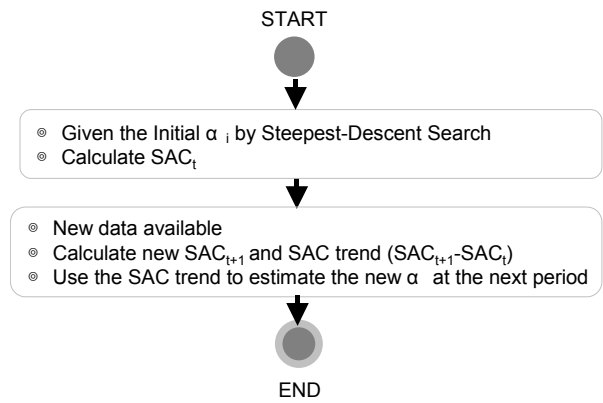


Fig. 7 Dynamic product mix estimate scheme

EVALUATION OF PROPOSED METHODOLOGY

In order to evaluate the proposed product-mix prediction scheme, we use proportion mean squared error (PMSE):

$$PMSE = \sum_{t=k+1}^{k+n} \sum_{i=1}^m (\hat{P}_{i,t} - P_{i,t})^2 / n$$

as the evaluation measure and compare the performance against two conventional proportion estimate methods: methods A and B.

In addition to the simulated DRAM demands, actually semiconductor demand data as shown in Fig. 6 is also tested.

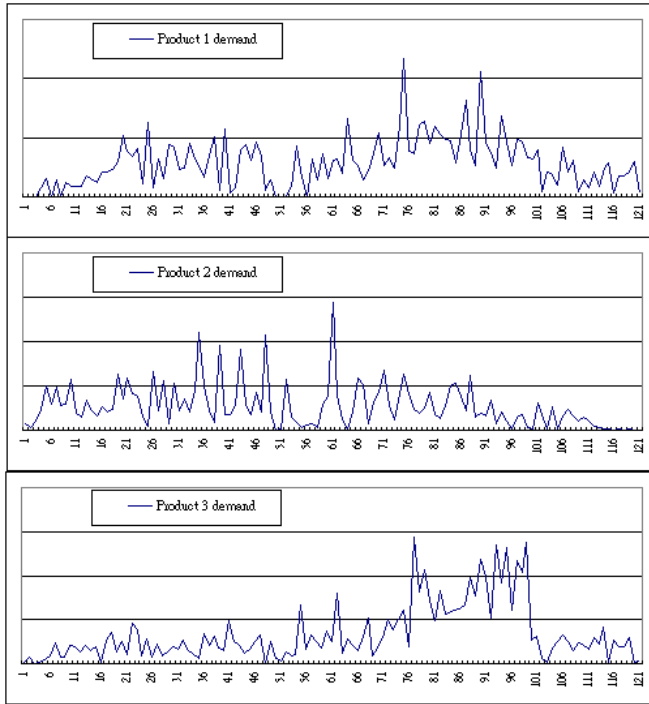


Fig. 8 Actual semiconductor demands

It is difficult to observe the relationship among the three individual demands from Fig. 8. However, the product-mix proportions (Fig. 9) show clearly the relationship among these three products.

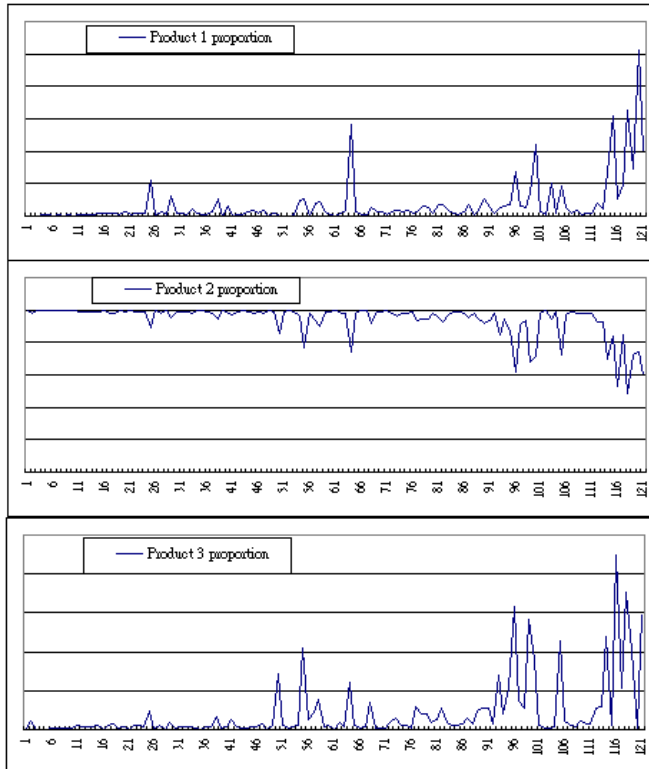


Fig. 9 Product-mix of actual semiconductor demands

The results (Table 1) show significant improvements by the proposed method (PLC Indicator Dynamic EWMA, PIDE) in both the simulated and actual demand cases.

Simulated Data

Conventional Method	Total PMSE
Method-A	0.072740
Method-B	0.064664
PIDE Method	Total PMSE
PIDE	0.001962

Actual Data

Conventional Method	Total PMSE
Method-A	0.009766
Method-B	0.011467
PIDE Method	Total PMSE
PIDE	0.007813

Table 1 Evaluation Results

ACKNOWLEDGEMENT

This research is supported in part by ISMT and SRC (879.001) and by NSC (NSC90-2218-E-002-046).

REFERENCE

- [1] Charles W. Gross and Jeffrey E. Sohl "Disaggregation Methods to Expedite Product Line Forecasting" Journal of Forecasting, Vol. 9, 233-254, 1990
- [2] Ward Cheney and David Kincaid, Numerical Mathematics and Computing, 4th edition, Brooks/Cole Publishing Company, 1999

AUTHOR BIOGRAPHY

Argon Chen is a faculty member of the Graduate Institute of Industrial Engineering at National Taiwan University (NTU). He received his Ph.D. degree in industrial engineering and earned his M.S. degrees in industrial engineering and in statistics from State University of New Jersey, Rutgers. Dr. Chen has been working closely with the semiconductor industry in Taiwan as a principle investigator of several research projects. In 2001, Dr. Chen was awarded a 3-year research project co-funded by Semiconductor Research Corporation (SRC) and International Sematech (ISMT). Dr. Chen published his research work mostly in *IEEE Transactions on Semiconductor Manufacturing* and *Technometrics*.

