Int J Adv Manuf Technol (2005) 26: 1429–1435
DOI 10.1007/s00170-004-2124-5

ORIGINAL ARTICLE

Muh-Cherng Wu · Kuo-Cheng Fu

An operator staffing and assignment model for foundry fabs

Received: 7 November 2003 / Accepted: 2 February 2004 / Published online: 12 January 2005

Abstract This paper presents a linear programming (LP) model for simultaneously solving an operator staffing problem in a semiconductor fab. The problem is to determine the staffing level and assign operators to the staffing positions. The semiconductor fab has the following characteristics: it imposes a no-lay-off policy, but allows overtime; keeps high-quality shop floor operations; aims to control tightly the staffing costs; and faces frequent changes of product mix. The operator staffing problem in such a fab was not addressed in previous literature. The proposed LP model aims to minimize the operator staffing cost. The LP solutions have been examined and were found to be insensitive to stochastic demands modeled by various simulation replicates.

Keywords Foundry fab · Operator assignment · Semiconductor manufacturing · Staffing

1 Introduction

Semiconductor wafer manufacturing is a machine-intensive industry. A typical wafer manufacturing facility (fab) includes about 300–500 machines. These machines are grouped into approximately 100 workstations. A workstation involves several machines that are functionally identical. On the shop floor, a team of operators is assigned to supervise a group of workstations, often called an area, a bay, a cluster or a work center [1].

Two main tasks performed by operators are the loading and unloading of wafers. When a machine is available for processing, operators must load wafers, then unload wafers when the machine has finished an operation. The machine becomes idle when no operator is present to perform loading or unloading tasks. The idleness of the machine due to the lack of an operator is rather expensive, which might result in the significant loss of capacity [2] and an increase in cycle time [3]. Using more operators can prevent machines from being idle, but it also increases staffing costs. Therefore, developing methods to make the operator staffing decision is important in determining how many operators should be placed at each work center to prevent machines from being idle while minimizing the staffing cost.

Many studies on the operator staffing decision have been published. Bonal et al. [1] develop a static operator staffing model using data obtained from the manufacturing execution system (MES). The static model calculates, on a weekly basis, the required operator time in a work center and the maximum available time an operator can provide. A staffing level is obtained by further considering the operators' availability and absenteeism. The static model assumes that the machine interference effect can be neglected. Machine interference occurs when the number of machines that simultaneously call for assistance exceeds the number of available operators, consequently causing production loss [4].

Some studies proposed an $M/M/s$ queuing model [5] to analyze the production loss caused by machine interference [6, 7] by modeling machines as customers and operators as servers. Meyersdorf et al. [6] proposed another queuing model to analyze the operator-machine-lot interference, in which wafer lots are taken as customers and operators as servers. The above studies based on queuing theory aimed to compute the minimum weekly demand of operators at which the production loss at a work center is tolerable.

The measurement of operator service time has not been thoroughly explored in the studies based on queuing. Kishimoto et al. [3] proposed a systematic method to estimate the service time of operators. They classified the activities of operators into three types. The first type, such as loading and unloading, is *move-based*, and the service time is proportional to the number of moves. A move refers to a machine's finishing of an operation; therefore, each move involves at least a loading and an unloading activity. The second type of activity is *time-based*; for example, a machine may need to be cleaned for two hours every two days.

M.-C. Wu (✉) · K.-C. Fu
Department of Industrial Engineering and Management,
National Chiao Tung University,
Hsin-Chu, Taiwan, R.O.C.
E-mail: mcwu@cc.nctu.edu.tw
Fax: +886-03-5720610

The third type is *administration-based*, and covers the service time for administration. A detailed analysis of the service time of operators led to the development of a heuristic method for reallocating the activities to operators. Such a reallocation, verified by a discrete event simulation model, outperforms the previous allocation of their testing fab.

In semiconductor wafer manufacturing, there are two types of business models: *foundry* and *integrated device manufacturing* (IDM). A typical IDM company manufactures only a few types of ICs (integrated circuits) that are designed by in-house. A foundry fab, by contrast, manufactures a large number types of ICs designed by many customers known as IC design houses. The product mix in a foundry fab may vary frequently, thus changing the staffing needs at each work center, and requiring that a decision be made regarding operator reallocation.

Most previous studies on fab operator staffing estimate a weekly staffing level at which the production loss or the increase in cycle time is acceptable. This implicitly advocates an idea that each work center should be run at a computed staffing level. *Staffing level* and *the number of staffing positions* are interchangeably used in this paper. However, three issues may arise when this idea is implemented in a foundry fab, which aims to tightly control its staffing costs, keep high-quality shop floor operation, and impose a no lay-off policy in an environment of a frequently changing product mix.

The first issue involves the fact that the staffing demands of operators at each work center may fluctuate daily. Staffing a work center at a particular level may cause some machines to be idle. One method of reducing machine idleness is by allowing operators work overtime. This implies that the computed demand of staffing level (or the number of staffing positions) may be reduced by overtime, yet this point has not been considered in previous studies.

The second issue concerns the fact that the available operators may be fewer than the computed demand. A foundry fab may insist on a no lay-off policy because the cost of lay-off is quite high in some non-Western societies. In addition to paying tangible compensation, a company that lays off employees may suffer some intangible but serious losses, such as of the morale of employees and the reputation of the company. A semiconductor company without a lay-off policy tends to hire operators conservatively. The supply of operators may thus be less than the computed demand when the product mix changes significantly. This issue, as well as the first issue outlined above, points out a decision problem: staffing position allocation; that is, determining how many staffing positions should be allocated to each work center.

The third issue is that the operators in a fab may have limited qualification for taking care of a work center. In a foundry fab, the shop floor operation is quite complex and dynamic. Operators must handle many dynamically changing events such as lot holding, route changing, and among-fab supporting, and familiarize with the reporting and coordination practices of handling exceptional cases in a particular work center. To ensure the quality of shop floor operation, a foundry fab may require that an operator receive rehearsal training if he or she has not been operating in a work center for a certain period of time (for example, six months). Without such training, operators might ignore some of the details of the operating practices, and subsequently cause adverse effects. The assignment of an operator to a work center may therefore include training cost. This points out another decision problem: operator assignment; that is, how to assign operators to the staffing positions such that the training costs are minimized.

This paper proposes a linear programming (LP) model to solve simultaneously the two aforementioned decision problems – staffing position allocation and operator assignment. The LP model is developed for cases in which the available staffing supply is less than the demand, but it can be extended to cases in which the available supply is greater than or equal to the demand. This study also assumes that an operator can only be assigned to one work center. This implies that each work center is an independent unit in the measurement of performance.

The rest of this paper is organized as follows: Sect. 2 presents a method for computing the staffing cost function of a work center, considering daily demand fluctuation and overtime. The staffing level for which the cost is minimum can be obtained by the staffing cost function. Section 3 describes the proposed LP model for cases in which the available staffing supply is less than the staffing level of minimum cost. Section 4 gives a numerical example to explain the proposed LP model and compares solutions in which the staffing demands are generated by simulation with various random seeds. Section 5 describes extensions of the LP model to accommodate cases in which the staffing supply is greater than the staffing level of minimum cost. Concluding remarks are made in Sect. 6.

2 Staffing cost at a work center

This section presents a method for estimating the cost of operators for a work center under various staffing levels in a foundry fab. Without loss of generality, the fab is assumed to implement the following staffing practices.

The shop floor of the fab is operated in a two-shifts-per-day schedule, with 12 hours in each shift. Subtracting the time for two meals and some other breaks in a shift, the availability of an operator is about 75%, or about 9 hours per shift. An operator typically works one shift per day for two consecutive days and then takes a break for the subsequent two days. Accordingly, the fab has four teams of operators, each of which works two shifts every four days.

The daily demand for operator time for the following few days can be estimated to reasonable accuracy using a discrete event simulation program, according to the answers given in an interview with a foundry fab conducted by the authors. Overtime is requested whenever the estimated daily demand exceeds the staffing level. An overtime request should be made on a shift basis; that is, a request of less than one shift of overtime is taken as a request for a whole.

The first step in computing the staffing cost function is to estimate the daily demand on operator time for a given master production schedule (MPS) in a quarter. Given the MPS, a discrete event simulation program coded by eM-Plant (http://www.tecnomatix.com/) is used to estimate the daily move of each work center. Assume that the operator service time for each type of move-based activity and time-based activity is known. The daily demand on operator time can be expressed as seen below:

$$x_j(t) = \frac{M_j(t)\, m_j}{60} + S_j(t)\,,$$

where

$x_j(t)$ = Demanded manpower in work center j on day t (man-hours)
$M_j(t)$ = Number of moves completed at work center j on day t
m_j = Service time per move at work center j (man-minutes/move)
$S_j(t)$ = Service time of time-based activities in work center j on day t (man-hours).

With $x_j(t)$, and the staffing policy as stated above, the operator cost of a work center is a function of the staffing levels:

$$C(p_j) = h(p_j\, 2L) + o \sum_{t=1}^{L} y_t \,\text{roundup}\left(\frac{x_j(t) - p_j\, 2 \times 12A_o}{12A_o}\right).$$

If $x_j(t) > (p_j\, 2 \times 12A_o)$, then $y_t = 1$, else $y_t = 0$, where

p_j = Number of staffing positions per shift at work center j
$C(p_j)$ = Total staffing costs of work center j with p_j positions (\$)
h = Normal time rate (\$/man-shift)
o = Overtime rate (\$/man-shift)
A_o = Availability of operators ($A_0 = 0.75$ in the fab of interest)
L = Total number of days in the decision time horizon.

The first term of the equation for $C(p_j)$ above denotes the normal time cost, where the time horizon includes $2L$ shifts; the second term represents the overtime cost, where $(x_j(t) - p_j\, 2 \times 12A_o)$ denotes the daily demand for overtime in man-hours. This value, when positive, should be represented in man-shifts (divided by $12A_o$) and rounded up to an integer.

Increasing the number of staffing positions (p_j) will increase the cost of normal time while decreasing the cost of overtime.

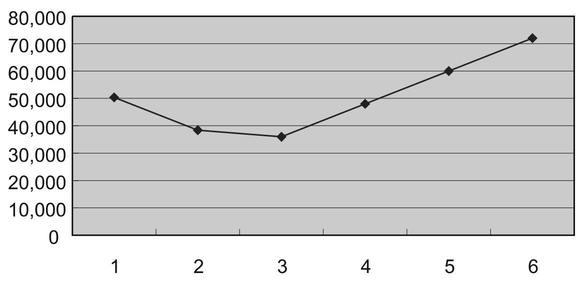

Fig. 1. Staffing cost function of a particular work center

The staffing cost function $C(p_j)$ is therefore a convex function with a global minimum point. By iteratively computing $C(p_j)$ at various staffing levels, we can easily determine a p_j^* which gives a minimum in staffing cost. Figure 1 shows a staffing cost function of a work center, where $p_j^* = 3$.

3 Model development

A linear programming model is proposed to solve the two decision problems – staffing position allocation and operator assignment. Let S be the number of operators available per shift, which, over the decision time horizon, cannot be increased according to the corporate staffing policy. Let $D_m = \sum_{j=1}^{J} p_j^*$ represent the minimum cost demand of staffing positions per shift, where J denotes the total number of work centers and p_j^* is the minimum cost demand of staffing positions at work center j. Without considering the constraints of operator supply and training requirements, D_m operators per shift should be supplied. However, when $S < D_m$, $D_m - S$ staffing positions should be eliminated to match S (the number of available operators). Therefore, the staffing position allocation problem is to determine which of the D_m staffing positions should be removed.

This study uses dummy operators to solve the problem of reducing the number of staffing positions. That is, $D_m - S$ dummy operators are created to make the operator supply equal D_m. The linear programming model is then used to assign each staffing position to an operator. The staffing positions assigned to the dummy operators are to be eliminated. The two decision problems, staffing position allocation and operator assignment, can therefore be integrated as an operator assignment problem.

Let w_{jk} represent the kth $(1 \le k \le p_j^*)$ staffing position at work center j $(1 \le j \le J)$. An encoding function $F(j, k) = s$ is designed such that each staffing position w_{jk} in the fab can be sequentially represented by q_s $(1 \le s \le D_m)$. A decoding function $F^{-1}(s) = j$ is defined to identify the work center to which q_s belongs. As stated, the operators in the fab may have various qualifications. All the operators are grouped such that each member of a group has the same qualifications. The dummy operators are placed in a single group. Let d represent the total number of groups and n_i represent the number of operators in group G_i $(1 \le i \le d)$. Then, $D_m = \sum_{i=1}^{d} n_i$.

The linear programming model formulates the two aforementioned decision problems as an operator assignment problem. Table 1 presents an example of the formulation, in which the first column represents the staffing positions at each work center, w_{jk} $(1 \le j \le J;\ 1 \le k \le p_j^*)$; the second column represents the staffing position q_s $(1 \le s \le D_m)$; and the first row represents the groups of operators G_i $(1 \le i \le d)$. In Table 1, $D_m = 10$ and $S = 8$; two dummy operators $(D_m - S)$ are thus created and placed in group G_6. Group G_1 includes one operator who is qualified only to take care of work centers 2 and 5. The operator must therefore be trained when he or she is assigned to supervise work centers 1, 3, or 4. Let C_{si} represent the cost of assigning a staffing

Table 1. Staffing position allocation and operator assignment matrix

w_{ij}	q_s	G_1	G_2	G_3	G_4	G_5	G_6 (Dummy)	Demand
w_{11}	q_1	T	0	T	T	T	$C(p_1^*-1)-C(p_1^*)$	1
w_{12}	q_2	T	0	T	T	T	$C(p_1^*-2)-C(p_1^*-1)$	1
w_{13}	q_3	T	0	T	T	T	M	1
w_{21}	q_4	0	T	T	T	0	$C(p_2^*-1)-C(p_2^*)$	1
w_{22}	q_5	0	T	T	T	0	M	1
w_{31}	q_6	T	T	T	T	0	$C(p_3^*-1)-C(p_3^*)$	1
w_{32}	q_7	T	T	T	T	0	M	1
w_{41}	q_8	T	0	0	0	T	$C(p_4^*-1)-C(p_4^*)$	1
w_{42}	q_9	T	0	0	0	T	M	1
w_{51}	q_{10}	0	T	0	0	T	M	1
Supply		1	2	3	1	1	2	10

position q_s to an operator in group G_i. For the non-dummy operator groups $(1 \le i \le d-1)$, $C_{si} = T$ if training is needed, while $C_{si} = 0$ otherwise. The matrix of C_{si} $(1 \le i \le d-1,\ 1 \le s \le D_m)$ therefore denotes the qualification profile of operators in the fab. Note that the example in Table 1 is simple; in fact, the training cost C_{si} is not necessarily a constant, when the operators have different capabilities in supporting a work center [9].

Assigning a dummy operator in G_d to a staffing position means that the position should be removed. The cost of removing the first staffing position from work center j can be represented by $C_{xd} = C(p_j^*-1) - C(p_j^*)$ where $x = F(j,1)$. Likewise, the cost of removing the second staffing position from work center j can be represented by $C_{xd} = C(p_j^*-2) - C(p_j^*-1)$, where $x = F(j,2)$. Notice that the formulation $C(p_j^*-2) - C(p_j^*-1)$ implies that the first staffing position has been removed. Therefore, there exists an inherent priority in removing the staffing positions from a work center j. That is, staffing position $w_{j,k+1}$ cannot be removed if w_{jk} has not been removed.

Moreover, the staffing position of a work center cannot be all removed. Such a removal implies that no operator will be assigned to the work center. The work center will subsequently become idle all the time. Removing the last staffing position from work center j is thus not allowed. The cost of doing so is represented by $C_{xd} = M$ where $x = F(j, p_j^*)$ and M is a very large positive number.

The operator assignment problem can be formulated as the following linear program. Let Q_{si} represent the binary decision variable; that is, $Q_{si} = 1$ if the staffing position q_s is assigned to an operator in group G_i, otherwise $Q_{si} = 0$.

$$\text{Minimize} \quad TC = \sum_{s=1}^{D_m} \sum_{i=1}^{d} C_{si} Q_{si} \tag{1}$$

subject to the following constraint sets:

$$\sum_{i=1}^{d} Q_{si} = 1 \qquad \text{for } 1 \le s \le D_m \tag{2}$$

$$\sum_{s=1}^{D_m} Q_{si} = n_i \qquad \text{for } 1 \le i \le d \tag{3}$$

$$Q_{sd} \ge Q_{(s+1)\,d} \qquad \text{for } 1 \le s \le D_m \text{ where } F^{-1}(s) = F^{-1}(s+1) \tag{4}$$

$$Q_{si} = 0 \text{ or } 1 \qquad \text{for } 1 \le s \le D_m\ ,\ 1 \le i \le d \tag{5}$$

The objective function Eq. 1 minimizes the total assignment costs, which are of two types. One is the cost of training C_{si} $(1 \le i \le d-1)$ which applies when real operators are assigned. The other is the cost of eliminating staffing position C_{sd} $(1 \le s \le D_m)$, which is incurred when dummy operators are assigned. Constraint set Eq. 2 specifies that each staffing position q_s be assigned to a single operator; and constraint set Eq. 3 requests that each operator in each group must be assigned to a staffing position.

Constraint set Eq. 4 denotes that the elimination of staffing positions in a work center should be performed in a predefined order. For example, in work center j, let q_s and q_{s+1} represent the first two staffing positions; that is, $F(j,1) = s$ and $F(j,2) = s+1$. If only one staffing position is to be removed from work center j, then q_s is removed. Skipping q_s and jumping to remove q_{s+1} is prohibited because $C_{(s+1)d}$ denotes the marginal cost of removing the second staffing position. Therefore, $Q_{sd} \ge Q_{(s+1)d}$ if q_s and q_{s+1} are in the same work center; that is, $F^{-1}(s) = F^{-1}(s+1)$.

The above LP formulation, an extension of a generalized assignment problem (GAP) [10], is distinguished by the inclusion of constraint set Eq. 4. Studies on the real-life applications of GAP include those of [11–17]. To the authors' knowledge, no previous GAP study has examined the staffing problem at a semiconductor fab with the features presented in this paper.

4 Examples and comparison

This section first describes a numerical example to explain the proposed methods for estimating staffing costs and assigning operators. Next, the operator assignment results are compared with values of daily demanded manpower generated by simulation using various random seeds.

4.1 Example

The fab in the example includes five work centers. The normal time rate is $h = \$\,1200$/shift and the overtime rate is $o = \$\,2400$/shift. The training cost for an operator is $T = \$\,3600$. The time horizon for the decision is five days.

The daily moves of each work center can be determined using a discrete event simulation program coded by eM-Plant (http://www.tecnomatix.com/). Table 2 presents the daily demanded operator hours at each work center given the service time rates for move-based and time-based activities. Accordingly, Table 3 presents the total cost over the decision time horizon for staffing each work center using different number of operators, where *NOP* denotes the number of staffing positions per shift.

Table 3 shows that $p_1^* = 3$, $p_2^* = 2$, $p_3^* = 2$, $p_4^* = 2$ and $p_5^* = 1$. Therefore, the demanded number of staffing positions

Table 2. Daily demand of operator hours at each work center

WS	Day 1	Day 2	Day 3	Day 4	Day 5
W1	43	37.4	45.8	40.2	43
W2	35	27	33	29	27
W3	36.2	33	29.8	29.8	33
W4	29.8	29.8	29.8	33.4	26.2
W5	17.8	17.8	20.2	22.6	20.2

Table 3. Staffing costs at each work center

W1		W2		W3		W4		W5	
NOP	Cost	NOP	Cost	NOP	Cost	NOP	Cost	NOP	Cost
1	50 400	1	31 200	1	38 400	1	33 600	1	19 200
2	38 400	2	24 000	2	26 400	2	24 000	2	24 000
3	36 000	3	36 000	3	36 000	3	36 000	3	36 000
4	48 000	4	48 000	4	48 000	4	48 000	4	48 000
5	60 000	5	60 000	5	60 000	5	60 000	5	60 000
6	72 000	6	72 000	6	72 000	6	72 000	6	72 000

Table 4. Result of operator assignment

	G_1	G_2	G_3	G_4	G_5	G_6	Demand
W1-1	3600	0	3600	3600	3600	2400	1
W1-2	3600	0	3600	3600	3600	12 000	1
W1-3	3600	0	3600	3600	3600	100 000	1
W2-1	0	3600	3600	3600	0	7200	1
W2-2	0	3600	3600	3600	0	100 000	1
W3-1	3600	3600	3600	3600	0	12 000	1
W3-2	3600	3600	3600	3600	0	100 000	1
W4-1	3600	0	0	0	3600	9600	1
W4-2	3600	0	0	0	3600	100 000	1
W5-1	0	3600	0	0	3600	100 000	1
Supply	1	2	3	1	1	2	10

per shift is $D_m = \sum_{j=1}^{5} p_j^* = 10$. Suppose that the supply of staffing operators per shift is $S = 8$. Then, two dummy operators must be created. Table 4 displays the cost matrix associated with operator assignment as well as the results obtained using the proprietary software package LINGO 5.0 (http://www.lindo.com/table/lingot.html/), where M (a large positive real number) is set to \$ 100 000 and G_6 represents the group of dummy operators. The total assignment cost is \$ 13 200 = \$ 2400 + \$ 7200 + \$ 3600. Table 4 reveals that two staffing positions are eliminated, one from work center 1 and the other from work center 2.

4.2 Comparison

The foregoing discussion assumes that the daily demanded manpower profile was obtained by simulation. One question may be raised: how sensitive is the operator assignment solution to the change of the daily demand profile, which is stochastic in the real world? This study therefore compares the assignment solutions

Table 5. Percentile distribution of ϱ_{ij} $(1 \le i \le 30,\ 1 \le j \le 30,\ i \neq j)$

10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
0	0	0	0	1.41	1.41	1.41	1.41	1.41	2.45

for various daily demand profiles, obtained by simulation with various random seeds.

A typical fab, data set 6 provided by the MASM Lab (http://www.eas.edu/~masmlab/home.htm), is used as a test bed to compare the operator assignment results. The fab includes 228 machines grouped into 104 workstations, which are grouped into 10 work centers. Nine types of products are produced in the fab.

In generating the daily demand profile, the simulation assumes that each machine is down at random, and that the downtime and repair time follow exponential distributions. Therefore, the operator assignment results associated with 30 simulation replicates are compared. In these experiments, the number of available operators (S) is 35.

Let $\overline{W_k} = [w_{k1}, w_{k2}, \ldots, w_{k10}]$ represent an operator assignment solution, where k denotes kth simulation replicate. An indicator $\varrho_{ij} = \left|\overline{W_i} - \overline{W_j}\right| = \sqrt{\sum_{k=1}^{10} (w_{ik} - w_{jk})^2}$ is defined to measure the distance between two solutions. Table 5 shows the percentile distribution of ϱ_{ij} $(1 \le i \le 30,\ 1 \le j \le 30,\ i \neq j)$, and indicates that 40% of ϱ_{ij} is zero, and 90% is less than or equal to 1.41. Notably, $\varrho_{ij} = 1.41 = \sqrt{2}$ means that the two solutions ($\overline{W_i}$ and $\overline{W_j}$) only differ in the assignment of one operator. This implies that the proposed method for operator assignment is insensitive to the stochastic variation of daily demand profile. That is, only one simulation replicate is required to solve the staffing problem with 90% confidence that the computed solution differs by no more than one operator from the best solution.

Note that the cases for $S = D_m$ are special cases of the above formulation, and can be easily solved.

5 Model extensions

The linear programming model described above can be extended to solve the operator assignment problems in cases in which $S > D_m$. For all D_m staffing positions, those in work center j are modeled as a group W_{j0}, where the number of staffing positions is p_j^*. When D_m positions are allocated, the fab may require that overtime be worked. As mentioned earlier, the fab imposes a no-lay-off policy. Therefore, when $S > D_m$, we may need to increase the number of staffing positions to match S. In doing so, the overtime cost may be reduced.

Let $V(p_j)$ represent the overtime cost function of work center j, which can be expressed as below, with reference to Sect. 2.

$$V(p_j) = o \sum_{t=1}^{L} y_t \text{ roundup} \left(\frac{x_j(t) - p_j 2 \times 12A_o}{12A_o} \right)$$

if $x_j(t) > (p_j 2 \times 12A_o)$, then $y_t = 1$, else $y_t = 0$.

At work center j with p_j^* staffing positions, the overtime cost is reduced by $V(p_j^*+1)-V(p_j^*)$ when one staffing position is added. Extra staffing positions can be added until no further reduction in overtime cost applies. That is, let e_j represent the total number of newly added positions at work center j. Then, $V(p_j^*+e_j+1)-V(p_j^*+e_j)=0$ and $V(p_j^*+e_j)-V(p_j^*+e_j-1)>0$. Let W_{jk} $(1 \le k \le e_j)$ represent the kth newly added staffing position at work center j. Remember that W_{jk} $(k=0)$ represents the p_j^* staffing positions at work center j. The total number of staffing positions at work center j is thus $p_j^*+e_j$.

Let D represent the total number of staffing positions including the original staffing level ($D_m = \sum_{j=1}^{J} p_j^*$) and the newly added staffing positions ($\sum_{j=1}^{J} e_j$); that is, $D = D_m + \sum_{j=1}^{J} e_j$. Note that D refers to the minimum cost staffing level without overtime. The operator assignment decision is to allocate S operators to D positions, which is to be addressed in the following two cases: $S > D > D_m$ and $D_m < S < D$.

5.1 $S > D > D_m$

When $S > D$, we need to create a dummy work center W_{d0} $(d = J+1)$, which includes $S-D$ staffing positions so that the supply and demand are equal. Table 6 shows an example, where $S = 15$, $D = 13$ and $D_m = 10$. The table includes $D - D_m = 3$ newly added staffing positions (W_{11}, W_{12}, W_{21}), and $S-D=2$ dummy positions in work center W_{60}.

The staffing positions (W_{jk}, $1 \le j \le J+1$, $0 \le k \le e_j$) are sequentially represented by b_s through an encoding function $s = Y(j,k)$. Let R_s denote $V(p_j^*+k)-V(p_j^*+k-1)$, where $s = Y(j,k)$, which is the marginal overtime cost reduction when the kth extra staffing position is added to worker center j. A decoding function $j = Y^{-1}(s)$ is defined accordingly.

In Table 6, the available operators are classified into groups (G_i) such that each operator in a group has the same qualifications. Let C_{si} represent the cost incurred when one staffing position in b_s is assigned to an operator in G_i. For the staffing group b_s representing W_{j0}, $C_{si} = T$ if the operator assignment needs a training; otherwise, $C_{si} = 0$. For the staffing group b_s representing W_{jk} $(k \ne 0)$, $C_{si} = T - R_s$ if the operator assignment requires a training; otherwise, $C_{si} = -R_s$. For the staffing group b_s representing dummy staffing positions (W_{d0}, $d = J+1$), $C_{si} = M$, where M is a large positive number. Assigning an operator to a dummy work center implies that no job will be assigned to the operator. The cost of such an assignment (M) should therefore be much higher than T and R_s.

Table 6. Assignment cost matrix for $S > D > D_m$

		G_1	G_2	G_3	G_4	Demand
W_{10}	b_1	T	0	T	T	3
W_{11}	b_2	$T-R_2$	$-R_2$	$T-R_2$	$T-R_2$	1
W_{12}	b_3	$T-R_3$	$-R_3$	$T-R_3$	$T-R_3$	1
W_{20}	b_4	0	T	T	T	2
W_{21}	b_5	$-R_5$	$T-R_5$	$T-R_5$	$T-R_5$	1
W_{30}	b_6	T	T	T	T	2
W_{40}	b_7	T	0	0	0	2
W_{50}	b_8	0	T	0	0	1
W_{60}	b_9	M	M	M	M	2
Supply		5	4	3	3	15

The linear program for $S > D > D_m$ is formulated below, where K represents the total number of operator groups and d represents the total number of staffing groups, including the dummy one. Let n_i represent the number of operators in operator group G_i, and m_s represent the number of positions in staffing group b_s. The term E_{si} $(1 \le s \le d; 1 \le i \le K)$ represents the cost of assigning an operator in G_i to staffing group b_s. The non-negative integer y_{si} $(1 \le s \le d; 1 \le i \le K)$ is a decision variable, which represents the number of operators in G_i assigned to staffing group b_s.

$$\text{Minimize} \quad TC = \sum_{s=1}^{d} \sum_{i=1}^{K} E_{si} y_{si} \tag{6}$$

such that

$$\sum_{s=1}^{d} y_{si} = n_i \qquad 1 \le i \le K \tag{7}$$

$$\sum_{i=1}^{K} y_{si} = m_s \qquad 1 \le s \le d \tag{8}$$

$$y_{si} \in Z \qquad 1 \le s \le d\,;\, 1 \le i \le K \tag{9}$$

The objective function in Eq. 6 models the total assignment costs. Constraint set Eq. 7 denotes that each operator in G_i should be assigned to a position, and set Eq. 8 specifies that each position in b_s should be assigned an operator. Note that this model is similar to the LP model presented in Sect. 3 except in the following two ways: first, the decision variable is not binary; second, constraint set Eq. 4 has been removed.

5.2 $D_m < S < D$

When $D_m < S < D$, a dummy group of operators must be created. Table 7 presents an example, where $S = 8$, $D = 10$ and $D_m = 7$. Note that the assignment cost for a dummy operator is M, and the other assignment costs are as discussed in Sect. 5.1.

Table 7. Assignment cost matrix for $D_m < S < D$

		G_1	G_2	G_3	G_K	Demand
W_{10}	b_1	T	0	T	M	3
W_{11}	b_2	$T-R_2$	$-R_2$	$T-R_2$	M	1
W_{12}	b_3	$T-R_3$	$-R_3$	$T-R_3$	M	1
W_{20}	b_4	0	T	T	M	2
W_{21}	b_5	$-R_5$	$T-R_5$	$T-R_5$	M	1
W_{30}	b_6	T	T	T	M	2
Supply		3	2	3	2	10

With the assignment cost matrix, a linear program model for $D_m < S < D$ can be formulated as follows: note that operator group K in the model denotes the dummy operator group.

$$\text{Minimize} \quad TC = \sum_{s=1}^{d} \sum_{i=1}^{K} E_{si} y_{si} \tag{10}$$

such that

$$\sum_{s=1}^{d} y_{si} = n_i \qquad 1 \le i \le K \tag{11}$$

$$\sum_{i=1}^{K} y_{si} = m_s \qquad 1 \le s \le d \tag{12}$$

$$y_{si} \in Z \qquad \text{for } 1 \le s \le d;\ 1 \le i \le K \tag{13}$$

$$y_{sK} \le y_{s+1,K} \qquad \text{for } Y^{-1}(s) = Y^{-1}(s+1) \tag{14}$$

This model is similar to the model presented in Sect. 5.1 except in that it includes constraint set Eq. 14, which is to ensure the appropriate assignments of dummy operators. With reference to Table 7, b_2 and b_3 are two newly added positions created in sequence. Therefore, b_2 cannot be assigned to a dummy operator while b_3 is assigned to a real operator.

The cases including $S = D$ or $D = D_m$ are special cases of the formulated models and can be easily solved.

6 Concluding remarks

This paper formulates the operator staffing problem faced by a foundry fab that aims to control tightly staffing costs, maintains a no lay-off policy, and ensures high-quality operation practices. A fab with these features cannot directly use previously described methods to staff work centers.

The staffing problem of such a fab is formulated as an LP model. The model, an extension of the model for solving the traditional GAP, is distinct in that it models some new constraints pertaining to the problem of interest. Much research on the application of GAP has been published. Yet, none has addressed the application of GAP to the problem of staffing a semiconductor fab.

The sensitivity of the LP solutions to the demand variations, modeled by various simulation replicates, has been examined. The study reveals that the operator assignment solution obtained by the proposed method is insensitive to stochastic variations of the daily demand profile. That is, only one simulation replicate is required to solve the staffing problem with 90% confidence that the computed solution differs by no more than one operator from the best solution.

The proposed LP method does not consider the machine interference effect. A future study will attempt to justify the proposed solutions, using queuing models or simulation models to evaluate the machine interference. That is, production loss due to machine interference must be evaluated for a work center that includes bottleneck machines. Minor adjustment of the proposed solution can therefore be made.

Acknowledgement The Taiwan Semiconductor Manufacturing Co. is acknowledged for providing this research problem and its shop floor management practices.

References

1. Bonal J, Fernandez M, Aparicio S, Oliva R, Rodriguez M, Rosendo M, Pablo C (2000) Operator Modeling. Proc Int Conf Model Anal Semicond Manuf, pp 104–106
2. Campbell GM, Diaby, M (2002) Development and evaluation of an assignment heuristic for allocating cross-trained workers. Eur J Oper Res 138:9–20
3. Campbell GM (1999) Cross-utilization of workers whose capabilities differ. Manage Sci 45:722–732
4. Cattrysse DG, Wassenhove LNV (1992) A survey of algorithms for the generalized assignment problem. Eur J Oper Res 60:260–272
5. Chen HN, Dabbas R (2002) Modeling staffing requirements within a semiconductor manufacturing environment. Proc IEEE/SEMI Adv Semicond Manuf Conf, pp 234–239
6. Creignou N, Hermann M (1996) Complexity of generalized satisfiability counting problems. Inf Comput 25:1–12
7. Fisher ML, Jaikumar R, Van Wassenhove L (1986) A multiplier adjustment method for the generalized assignment problem. Manage Sci 32:1095–1103
8. Fisher ML, Jaihumar R (1981) A generalized assignment heuristic for vehicle routing. Networks 11:109–124
9. de Fluiter B, Aarts EHL, Korst JHM, Verhaegh WFJ, van der Werf A (1996) The complexity of generalized retiming problems. IEEE Trans on Computer-aided design of integrated circuits and systems 15: 1340–1353
10. Kishimoto M, Ozawa K, Watanabe K, Martin D (2001) Optimized operations by extended X-factor theory including unit hours concept. IEEE Trans Semicond Manuf 14:187–195
11. Meyersdorf D, Biron O, Ozelkan C (1997) Staffing analysis tool for operator-machine-lot interference in semiconductor manufacturing. Proc IEEE/SEMI Adv Semicond Manuf Conf, pp 335–340
12. Mazzola JB, Neebe AW, Dunn CVR (1989) Production planning of a flexible manufacturing system in a material requirements planning environment. Int J Flexible Manuf Syst 1:115–142
13. Pollitt C, Matthews J (1998) Quantifying capacity loss associated with staffing in a semiconductor manufacturing line. Proc IEEE/SEMI Adv Semicond Manuf Conf, pp 133–137
14. Raviv A (1995) Applications of queuing theory and simulation to staffing in the semiconductor clean room environment. Int Symp Semicond Manuf, pp 252–255
15. Stecke K, Aronson JE (1985) Review of operator/machine interference models. Int J Prod Res 23:129–151
16. Winston WL (1991) Operations research applications and algorithms, 2nd edn. PWS-Kent, Boston
17. Shtub A, Kogan K (1998) Capacity planning by the dynamic multi-resource generalized assignment problem. Eur J Oper Res 105:91–99