

Scheduling for Wafer Fabrication System Based on Neuro-Dynamic Programming

WANG Ying, LIU Jie

Department of Automation, Xiamen University, Xiamen, China

e-mail: Wangying@xmu.edu.cn

Abstract: Scheduling of wafer fabrication system with stochastic machine failures time, repair time and processing time is studied by Neuro-Dynamic Programming (NDP) in this paper. Proper dynamic programming model are constructed to try to obtain the scheduling model of wafer fabrication system, on which NDP can be applied successfully and make satisfactory scheduling decisions. States set and scheduling set are constructed to describe the manufacturing states accurately and Markovly. Scheduling of wafer fabrication system is semi-Markov decision process, so dynamic programming model with closed-release policy is deduced after discretization of Markov and stochastic shortest path conversion from average-cost problem. Key features are derived by neglecting non-bottleneck stations, and then scheduling policy can be obtained by approximation of optimization cost function with simulation. Technology Research Center Silicon fab (TRC) model of HP is scheduled to illustrate the validity of the method.

Keywords: Neuro-dynamic Programming; Wafer fabrication; Scheduling; Feature simplification

基于神经元动态规划的半导体生产线调度

王颖, 刘杰

厦门大学自动化系, 厦门, 中国, 361005

e-mail: wangying@xmu.edu.cn

【摘 要】采用神经元动态规划针对机器故障及维修时间、加工时间随机的半导体生产线进行调度研究, 论文从建立合理的动态规划模型入手, 尝试构建半导体生产线的调度模型, 有效的解决了神经元动态规划无法直接有效的应用在半导体生产线调度中的问题。文章建立合理的状态变量和调度集, 以准确描述系统的生产状况, 并保证其 Markov 性; 针对半导体生产调度为半 Markov 决策过程的特性, 通过 Markov 离散化和平均代价问题的随机最短路径化, 推导出封闭式投料策略下的动态规划模型; 通过简化掉非瓶颈工作站来抽取关键特征变量, 进而获得近似最优函数进行调度, 最后以 HP 公司提出的 Technology Research Center Silicon fab (TRC) 模型为例进行调度寻优。

【关键词】神经元动态规划; 半导体生产线; 调度; 特征简化

1 引言

半导体制造系统被认为是当今最为复杂的制造系统之一, 其特有的“重入性”使其既不同于 Job-shop 问题也不同于 Flow-shop 问题, 被 Kumar 抽象为“可重入生产系统 (Re-entrant lines)”^[1], 是典型的离散事件动态系统^[2, 3]。半导体制造系统的生产过程调度问题被证明是 NP 完全问题, 随着系统规模的增大, 带来“维数灾”问题, 采用传统的规划求解方法难以奏效。此外, 大规模、不确定环境下的复杂生产过程

动态调度问题的研究, 也超越了传统调度理论的研究模式, 因此求解半导体生产线的优化调度算法一直是学术界的研究热点问题。目前, 国内外常用的半导体生产线调度算法包括运筹学方法、启发式方法、人工智能方法、计算智能方法以及群体智能方法^[4, 5]。

神经元动态规划 (Neuro-Dynamic Programming, NDP) 由 Bertsekas 和 Tsitsiklis^[6]于 1995 年提出, 它利用函数近似简化对系统状态集的搜索, 求取次优或者局部优的决策方案, 为克服“维数灾”问题提供了一条可行的道路。文献[7]尝试将神经元动态规划应用于可重入生产系统的调度寻优中, 但由于其应用过程

本文受福建省自然科学基金(2009J01304)资助

中存在以下几个问题：①状态的设计无法准确描述复杂的半导体生产线的生产状况；②忽视了半导体生产线为半 Markov 决策过程无法直接运用神经元动态规划方法的问题；③真正的半导体生产线状态复杂，将状态的每个分量都作为神经网络的特征变量明显是不合适也无法实现的；使得[7]中神经元动态规划的应用效果并不明显，且只能应用到二站四工序无故障的理想模型中，无法调度更加复杂的实际半导体生产过程。

本文从建立合理的动态规划模型入手，尝试构建半导体生产线的调度模型，在此模型上能顺利地应用神经元动态规划解决半导体生产的调度问题。文章针对具有机器故障及机器维修时间随机、加工时间随机的半导体生产线建立合理的状态变量和调度集，以准确描述系统的生产状况，且保证其 Markov 性，分析状态转移概率，针对半导体生产调度过程为半 Markov 决策过程的特性，通过 Markov 离散化和平均代价问题的随机最短路径化，推导出封闭式投料策略下的动态规划模型，并通过简化非瓶颈工作站来抽取反映系统主要矛盾的特征变量，获得合理的神经网络结构，从而获得近似的调度优化函数，最后以 HP 公司提出的 Technology Research Center Silicon fab (TRC) 模型为例进行调度寻优。

2 系统描述

关于半导体生产线的描述见文献[8]，进一步假定：工作站 W_i 中的机器故障的间隔时间、维修机器所需的时间分别服从均值为 $MTBF_i$ 和 $MTTR_i$ 的指数分布，机器故障包括无法预测的机器故障、事先安排的机器维护和调试。机器的故障和维修都是非清空的，即可以在工件加工未完成时发生故障，维修排除故障后工件重新开始加工。

考虑满足以下描述的工作站称为瓶颈工作站，其制约了系统平均输出率。在我们所描述的系统内，每个工件在工作站 W_i 上的平均加工时间为

$$T_j = \sum_{W(k)=W_j} \frac{1}{\mu_k}, \quad T'_j = \frac{T_j}{S_j} \text{ 为工作站 } W_j \text{ 完成单位工件各}$$

工序的总平均加工时间，则 $\frac{MTBF_j}{T'_j}$ 为机器一次正常运行能加工的工件个数，

$$n_j = \frac{MTBF_j}{T'_j \times (MTBF_j + MTTR_j)} \text{ 为}$$

考虑故障及维修时间，工作站 W_j 的最大可能平均输出率。若 $n_k = \min_{1 \leq j \leq L} \{n_j\}$ ，则称工作站 W_k 是系统的一个瓶

颈工作站。

2.1 状态变量和调度集描述

欲将系统描述为动态规划过程，其状态必须满足 Markov 性，因此系统状态表示为：

$$x = \left(n_{1w}, \dots, n_{Mw}, b_{11}, J_{11}, b_{12}, J_{12}, \dots, b_{1S_1}, J_{1S_1}, \dots, b_{L1}, J_{L1}, b_{L2}, J_{L2}, \dots, b_{LS_L}, J_{LS_L} \right) \quad (2.2)$$

半导体生产线中，按照系统当前是否需要调度，可以将状态分成两类。本文称其为待调度状态和非待调度状态，分别用 VS 和 TS 表示其状态集合。待调度状态表示系统正在等待调度，由于工件传输时间为 0，产生调度的时间为 0，故待调度状态的逗留时间也为 0。可数学描述成存在某个组合 (i, j) ，满足 $b_{ij} = 0, J_{ij} = 0$ 且 $n_{kw} \neq 0, W(k) = W_i$ ，则当前状态为待调度状态。非待调度状态是无须调度的状态，此时所有的机器要么在加工某个工件，要么出于某种原因不能加工工件。待调度状态是一种瞬时状态，在该状态下采取不同的决策，会转移到不同的非待调度状态，而非待调度状态是一种持续状态，它将持续到新工件进入、某个工件完成加工、机器产生故障或完成维修，以某种概率自动转移到待调度状态，等待下一步的调度。因此系统的优化调度问题可以描述为分阶段决策问题：为待调度状态集 VS 中的每个状态寻找合适的调度策略，以最优化某个性能指标。调度一旦完成，系统就进入非待调度状态。系统在待调度状态和非待调度状态的交替中演化。相应的调度决策表示为 $u = (y_1, y_2, \dots, y_M)$ ，决策下一步加工第 i 个工序的工件则 $y_i = 1$ ，否则 $y_i = 0$ 。

当系统的待调度状态满足式 2.3，则称系统处于阻塞状态，该状态与瓶颈工作站相关。

$$\begin{aligned} \exists i, j, b_{ij} = 0, J_{ij} = 0 \\ \text{且 } \sum_{k, W(j_k)=W_i} n_{kw} = 0, W_i \text{ 为瓶颈工作站} \end{aligned} \quad (2.3)$$

2.2 状态转移概率

系统生产过程中，从待调度状态到非待调度状态的转换是确定的，从非待调度状态到待调度状态的转换则是随机的，可推导出系统待调度状态之间的一步转移概率：

$U(x_i)$ 表示对待调度状态 x_i ，所有将系统转移到非待调度状态的调度集， $\psi(x_i, u)$ 表示在待调度状态 x_i 采取调度 $u \in U(x_i)$ 所致的非待调度状态中所有正在加工工序的集合。当 $J_{ij} \neq 0, b_{ij} = 1$ 时， $J_{ij} \notin \psi(x_i, u)$ 。

$S(l) \in VS$ 表示完成工序 $l, l \in \psi(x_i, u)$ 后所转移到的待调度状态。

由上可知, 半导体生产过程可视为其待调度状态集 VS 上的半 Markov 决策过程, 即该过程受一系列控

$$p_{x_i x_j}(u) = \begin{cases} \frac{\sum_{l \in \psi(x_i, u), S(l)=x_j} \mu_l}{\sum_{k \in \psi(x_i, u)} \mu_k + \sum_{k \in q} (1-b_{kq}) \frac{1}{MTBF_k} + \sum_{k \in q} b_{kq} \frac{1}{MTTR_k}} & \text{下一状态为完成工序的概率} \\ \frac{1}{MTBF_i} & \text{下一状态为第 } i \text{ 个工作站上的某台机器发生故障的概率} \\ \frac{\sum_{k \in \psi(x_i, u)} \mu_k + \sum_{k \in q} (1-b_{kq}) \frac{1}{MTBF_k} + \sum_{k \in q} b_{kq} \frac{1}{MTTR_k}}{\sum_{k \in \psi(x_i, u)} \mu_k + \sum_{k \in q} (1-b_{kq}) \frac{1}{MTBF_k} + \sum_{k \in q} b_{kq} \frac{1}{MTTR_k}} & \text{下一状态为第 } i \text{ 个工作站上某台正在维修的机器完成维修的概率} \end{cases}$$

制决策驱动, 其状态转移规律和采取的控制决策相互作用, 决定了系统的演化过程在每个状态的逗留时间是服从一般分布的随机变量, 其间所含的 Markov 链是互达的^[9]。

3 动态规划模型的推导

神经元动态规划方法需要运用在动态规划的模型框架上, 因此如何合理有效的将半导体生产过程表达为动态规划模型则尤为重要。

3.1 半 Markov 决策过程的离散化

由于半导体生产过程为半 Markov 决策过程, 为克服其仿真优化上的困难, 本节将其转化为离散的 Markov 链。

在封闭式投料策略下, 可将全部工作时间内瓶颈工作站内所有机器同时工作所占的比例 η 作为性能指标, 作为平均输出率 ξ 的有效替代^[10]。因此采用瓶颈工作站内所有机器同时工作所占的比例 η 作为等价性能指标, 推导动态规划模型。

设系统在策略 μ 下最终到达稳态分布的非待调度状态数为 H , 系统为连续时间 Markov 过程 Y , 记 Y 在策略 μ 下的稳态分布为 p_μ 。对连续时间 Markov 过程 Y 取性能函数:

$$f(y_i) = \begin{cases} 1 & y_i \notin B \\ 0 & y_i \in B \end{cases} \quad (3.1)$$

其中 B 表示系统所有的阻塞状态的集合, 当待调度状态 $y_i \in B$ 时, 表示对应的调度集为空, 即系统中至少存在有一个瓶颈工作站中的某台机器处于无工件

可以加工的状态。

在策略 μ 下, 瓶颈工作站内所有机器同时工作所占的比例:

$$\eta_\mu = \sum_{i=0}^H p_\mu(y_i) f(y_i) \quad (3.2)$$

过程 Y 中, 当待调度状态 y_i 采取策略 μ 时, 得到对应的非待调度状态的平均逗留时间记为 $m_\mu(y_i)$, $\{T_l; l \geq 0\}$ 为系统的状态转移时刻, $X = \{X_l; l \geq 0\}$ 为 Y 的嵌入 Markov 链, 其中 $X_l = Y_{T_l+0}$, 由系统中的待调度状态组成。设 X 在策略 μ 下的稳态分布为 π_μ , 并记

$$\sigma_\mu = \sum_{i=1}^H \pi_\mu(x_i) m_\mu(x_i) \quad (3.3)$$

取嵌入链 X 的性能函数如式 3.4 所示, 则可得到式 3.5, 因此 X 的性能指标为式 3.6:

$$\tilde{f}_\mu(x_i) = \begin{cases} m_\mu(x_i) / \sigma_\mu & x_i \notin B \\ 0 & x_i \in B \end{cases} \quad (3.4)$$

$$\tilde{f}_\mu = \frac{\text{diag}(m_\mu(x_1), \dots, m_\mu(x_H))}{\sigma_\mu} \cdot f \quad (3.5)$$

$$\tilde{\eta}_\mu = \sum_{i=1}^H \pi_\mu(x_i) \tilde{f}_\mu(x_i) \quad (3.6)$$

由于对于半 Markov 过程及其对应的嵌入 Markov 链, 有式 3.7^[11]

$$p_\mu \text{diag}(m_\mu^{-1}(x_1, u), \dots, m_\mu^{-1}(x_H, u)) = \frac{\pi_\mu I}{\sigma_\mu} \quad (3.7)$$

由式 3.2、3.5、3.6、3.7 可得:

$$\tilde{\eta}_\mu = \eta_\mu \quad (3.8)$$

则连续 Markov 决策过程 Y 与其嵌入式 Markov 链 X 具有相同的性能指标。将 $\tilde{\eta}_\mu$ 写成平均代价问题的动态规划形式, 记一步转移代价为:

$$g(x_i, \mu(x_i), x_j) = \begin{cases} -m_\mu(x_i) / \sigma_\mu & x_i \notin B \\ 0 & x_i \in B \end{cases} \quad (3.9)$$

此时

$$\begin{aligned} J_\mu(x) &= \lim_{K \rightarrow \infty} \frac{1}{K} E \left(\sum_{k=1}^{K-1} g(x_k, \mu(x_k), x_{k+1}) \mid x_0 = x \right) \\ &= \frac{1}{\sigma_\mu} \lim_{K \rightarrow \infty} \frac{1}{K} E \left(\sum_{k=1}^{K-1} \sigma_\mu g(x_k, \mu(x_k), x_{k+1}) \mid x_0 = x \right) \end{aligned} \quad (3.10)$$

在封闭式投料策略下, 当工件数目趋于无穷时, 各非空闲策略都有 $\lambda_\mu \rightarrow \lambda^*$, σ_μ 可近似为一常数^[12], 问题转化为:

$$J_\mu(x) = \lim_{K \rightarrow \infty} \frac{1}{K} E \left(\sum_{k=1}^{K-1} \sigma_\mu g(x_k, \mu(x_k), x_{k+1}) \mid x_0 = x \right) \quad (3.11)$$

3.2 平均代价问题转化为最短路径问题

平均输出率和全机同时工作比例对应平均代价问题，在优化实现上有许多困难，理论体系不够完备，比较而言随机最短路径问题实现的可操作性更强，收敛性等理论体系更加完备^[6]，因此考虑将平均代价问题转化为随机最短路径问题。

构造一个 Markov 链 X' ， X' 包含 X 的全部状态和一个人工状态 $\bar{s}$ ， B 表示阻塞状态的集合， X 和 X' 中的转移概率记为 $p_{x_i x_j}^X(u)$ 和 $p_{x_i x_j}^{X'}(u)$ 。

在 X 中，当状态 $x_j \notin B$ ，则 $p_{x_i x_j}^{X'}(u) = p_{x_i x_j}^X(u)$

在 X 中，当状态 $s \in B$ ，则 $p_{x_i s}^{X'}(u) = 0, p_{x_i \bar{s}}^{X'}(u) = p_{x_i s}^X(u)$

改造前后的 Markov 链是等价的^[6]，当系统从状态 s 出发，一旦到达状态 $\bar{s}$ 则链即终止，不再转入其他状态。因此针对平均代价的 Markov 链 X 中的一条无穷长轨迹就对应着 X' 中无穷条起始于阻塞状态并终止于阻塞状态的轨迹，则 X 中的平均代价优化问题就等价于在 X' 中，力求找到一个策略 μ 最小化平均每步代价：

$$\lambda_\mu = \frac{C_{s_i s_j}^\mu}{N_{s_i s_j}^\mu} \quad (3.12)$$

其中 s_i, s_j 为任意阻塞状态， $C_{s_i s_j}^\mu$ 是 μ 策略下从状态 s_i 出发首次到达一个阻塞状态 s_j 付出的代价的数学期望，而 $N_{s_i s_j}^\mu$ 是从状态 s_i 到状态 s_j 的步数的数学期望。记 λ^* 为平均每步最优代价， $C_{s_i s_j}^\mu - N_{s_i s_j}^\mu \lambda^*$ 即表示从 s_i 出发，当一步转移代价为式 3.13 时采取策略 μ 的随机最短路径问题的期望代价。

$$\begin{cases} -m_\mu(x_i) - \lambda^* & x_i \notin B \\ -\lambda^* & x_i \in B \end{cases} \quad (3.13)$$

记 $h^*(x_i)$ 为从各 $x_i \notin B$ 出发、一步转移代价为式 3.13 的随机最短路径问题的最优解，于是存在唯一的 $h^*(x_i)$ 满足 Bellman 方程：

$$h^*(x_i) = \min_{u \in U(x_i)} \left[\sum_{x_j=1}^n p_{x_i x_j}(u) (g(x_i, u, x_j) - \lambda^* + h^*(x_j)) \right] \quad (3.14)$$

状态 $x_i, x_j \in VS$ 为系统的待调度状态。系统的一步转移代价 $g(x_i, u, x_j)$ 为：

$$g(x_i, u, x_j) = \begin{cases} -m_\mu(x_i) = -\frac{1}{\sum_{k \in \psi(x_i, u)} \mu_k} & x_i \notin B \\ 0 & x_i \in B \end{cases} \quad (3.15)$$

式 3.14 既是最短路径问题又是平均代价问题的 Bellman 方程， $h^*(x_i)$ 表示从状态 x_i 出发，第一次到达状态 $s \in B$ 的期望代价与当每步代价为 λ^* 时产生总代价的最小差值。

4 面向特征抽取的模型简化

由于实际的半导体生产线工序繁多，涉及到的状态维数巨大，在选择特征时若考虑系统的每个状态分量将使得神经网络过于庞大。本文认为系统中的瓶颈工作站是影响系统调度的主要矛盾，因此选取特征时主要关注瓶颈工作站的工作状态，同时兼顾其他工作站上工件个数对调度的影响。本文针对原系统进行仿真，将仿真获得的状态通过原系统状态简化算法加以简化，将非瓶颈工作站的状态整合到瓶颈工作站上，从中提取出相应的状态作为近似函数的特征量，获取相应的调度决策后通过调度策略的还原算法将其还原应用于原系统中。系统原始状态如式 4.1 所示，通过式 4.1 转换为 $x = (n_{1w}, \dots, n_{Mw}, n_{1p}, \dots, n_{Mp})$ 的形式以便于特征抽取，其中 n_{kp} 表示正处于第 k 道加工工序的工件的个数。

$$n_{kp} = \sum_{j=k}^M 1, k=1, \dots, M \quad (4.1)$$

将所有瓶颈工作站上的工序按其加工先后为序描述为 $I_1, I_2, \dots, I_D$ ， D 为瓶颈工作站上加工的工序的总个数，则简化后的系统状态描述为 $z = (m_{I_1w}, m_{I_1p}, m_{I_2w}, m_{I_2p}, \dots, m_{I_Dw}, m_{I_Dp})$ ，相应的调度策略描述为 $u = (y_{I_1}, y_{I_2}, \dots, y_{I_D})$ ，决策下一步加工第 I_i 个工序的工件则 $y_{I_i} = 1$ ，否则 $y_{I_i} = 0$ 。

原系统的状态简化：

$$m_{I_1w} = n_{(I_D+1)w} + n_{(I_D+1)p} + n_{(I_D+2)w} + n_{(I_D+2)p} + \dots + n_{Mw} + n_{Mp} + n_{1w} + n_{1p} + n_{2w} + n_{2p} + \dots + n_{(I_1-1)w} + n_{(I_1-1)p} + n_{I_1w}$$

$$m_{I_1p} = n_{I_1p}$$

$$m_{I_iw} = n_{(I_{(i-1)}+1)w} + n_{(I_{(i-1)}+1)p} + n_{(I_{(i-1)}+2)w} + n_{(I_{(i-1)}+2)p} + \dots + n_{(I_i-1)w} + n_{(I_i-1)p} + n_{I_iw}$$

$$m_{I_ip} = n_{I_ip} \quad i=2 \dots D$$

调度：针对简化系统获得调度规则 $u = (y_{I_1}, y_{I_2}, \dots, y_{I_D})$ ，当决策 $y_{I_i} = 1$ ，则 $m_{I_iw} = m_{I_iw} - 1, m_{I_ip} = m_{I_ip} + 1$

原系统相应状态变化：

如果 $n_{I_iw} \neq 0$ ，则 $n_{I_iw} = n_{I_iw} - 1, n_{I_ip} = n_{I_ip} + 1$

如果 $n_{I_{jw}} = 0$ 且 $\exists n_{I_{jw}} \neq 0$ ($I_j, W(I_j) = W(I_i)$),
则任选 I_j , 令 $n_{I_{jw}} = n_{I_{jw}} - 1$, $n_{I_{jp}} = n_{I_{jp}} + 1$
如果 $n_{I_{jw}} = 0$ 且 $n_{I_{jw}} = 0$ (for all $I_j, W(I_j) = W(I_i)$),
则机器处于空闲状态

本文从以下 4 个方面构建近似函数的特征向量:

(1) 简化后状态对应各工序正在加工、等待加工的工件的个数, 即 $m_{I_{jw}}, m_{I_{jp}}$ 分量。

(2) 简化后各分量 $m_{I_{jw}}, m_{I_{jp}}$ 的二次项, 包括自身二次项及交叉二次项。

(3) 从系统负载平衡的角度出发, 将各瓶颈工作站负载两两相比, 将比值作为特征分量:

$$\frac{(m_{I_{jw}} + m_{I_{jp}})\mu_i}{(m_{I_{jw}} + m_{I_{jp}})\mu_j} \quad (4.2)$$

(4) 由于启发式策略在某种程度上已经抓住优化策略的主要特点, 因此在此基础上合理训练参数向量, 能获取更优的策略。针对启发式策略 $\mu_k, k=1, \dots, K$, 通过仿真训练获取策略 μ_k 对应的平均代价的近似函数形式 $\hat{\lambda}^{\mu_k}$, 将其作为特征加以合适的权重逼近 $h^*(x_i)$, 则近似代价函数写成:

$$\hat{h}(x_i, r) = w_0(x_i, r_0) + \sum_{k=1}^K w_k(x_i, r_k) \tilde{\lambda}^{\mu_k} \quad (4.3)$$

$$w_k(x_i, r_k) = r_k(0) + \sum_{m=1}^M r_k(m) f_m(x_i) \quad (4.4)$$

则有式 4.5, 该式中可看出采用了 $f_m(x_i)$, $\hat{\lambda}^{\mu_k}$ 和 $f_m(x_i) \hat{\lambda}^{\mu_k}$ 作为特征量:

$$\begin{aligned} \hat{h}(x_i, r) = & r_0(0) + \sum_{m=1}^M r_0(m) f_m(x_i) + \sum_{k=1}^K r_k(0) \hat{\lambda}^{\mu_k} \\ & + \sum_{m=1}^M \sum_{k=1}^K r_k(m) f_m(x_i) \hat{\lambda}^{\mu_k} \end{aligned} \quad (4.5)$$

在训练过程中应注意特征各分量的归一化, 防止由于各特征分量所占的初始权重差别过大而严重影响神经网络学习效果。接下来的训练过程与一般神经网络的训练无异, 在此不再赘述。

因此, 通过离线训练获得近似代价函数, 在线使用该函数调度寻优, 从而获得生产线上满意的调度策略。

5 TRC 模型调度寻优

采用上文提出的方法对 HP 公司提出的半导体生产标准问题 TRC 模型进行调度仿真求解。该模型的绝大部分参数均采集自实际系统 Hewlett-Packard Technology Research Center Silicon fab, 具有 24 个工

作站 172 道工序, 对该模型的详细描述参见文献[13], 该模型提出后被视为半导体生产线的重要标准问题, 被许多学者作为验证其提出方案有效性的重要例证[14-17]。该系统有 FAB1, FAB2, FAB3 三种零件加工系统, 对 FAB1, 14 号工作站是瓶颈; 对 FAB2, 第 14 号和第 24 号工作站是瓶颈; 对于 FAB3, 工作站 14、19、21、24 为瓶颈工作站。采用封闭式投料策略 $N = 20$ 不变。系统的初态选择为 $[N, 0, \dots, 0, 0, 0, 0]$, 参数 $\lambda = 0.5$, 学习参数

$$\eta_r = \eta_\lambda = \frac{0.05}{(1 + \text{step})}$$

对每条状态轨迹学习后 step 递增 1。训练获得策略迭代走势如图 1 所示, 与其他调度策略比较如表 1 所示。比较看出采用神经元动态规划学习获得的策略性能明显优于原有的许多策略。虽然获得逼近函数需要大量时间, 但得到逼近函数后, 能获取近实时的调度策略。

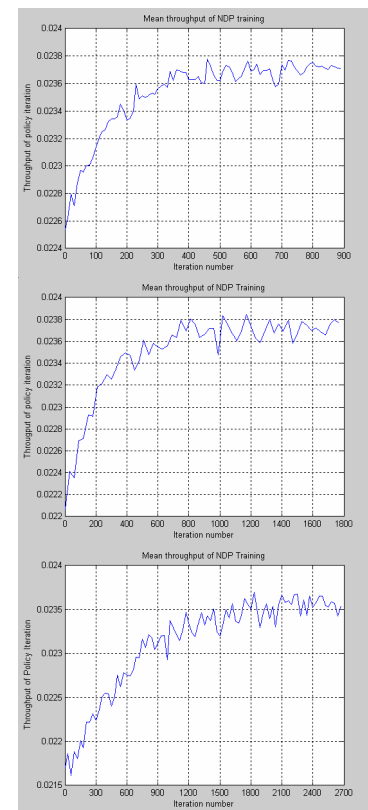


Figure 1. Scheduling policy iterations of FAB1, FAB2, FAB3

图 1. FAB1、FAB2、FAB3 调度策略迭代性能走势图

6 结论

本文采用神经元动态规划对半导体生产过程进行

调度研究, 通过合理的构建和分析转换推导出有效的动态规划模型, 在此模型上能顺利地应用神经元动态规划有效的解决半导体生产的调度问题, 通过对 HP

Table 1. Policies' performance of FAB1, FAB2, FAB3
表 1. FAB1、FAB2、FAB3 生产系统各策略下的性能表

Policy	Mean throughput		
	FAB1	FAB2	FAB3
FIFO	0.02242	0.02245	0.02158
SRPT	0.02260	0.02265	0.02156
FGCA	0.02262		
LWNQ/M	0.02250	0.02212	0.02157
FIFO+	0.02235	0.02287	0.02188
SRPT+	0.02254	0.02293	0.02193
CYCLIC	0.02251		
LTVN	0.02232		
STNV	0.02161		
MODD	0.02291		
PBDR	0.02300		
FSMCT	0.02341	0.02340	0.02331
FSVCT	0.02324	0.02291	0.02281
FGCA/IMP		0.02229	
M1-M2		0.02268	
NDP	0.02370	0.02373	0.02360

公司的 TRC 模型的调度寻优来验证该方法的有效性, 并获得较好的效果。

References (参考文献)

- [1] Kumar, P.R., Re-entrant lines. Queueing Systems: Theory and Applications[J], Special Issue on Queueing Networks, 1993. 13, P87-110.
- [2] Kim Y D, S.S.O., Choi B, Simplification methods for accelerating simulation-based real-time scheduling in a semiconductor wafer fabrication facility[J]. IEEE Transactions on Semiconductor Manufacturing, 2003. 16(2), P290-298.
- [3] Hsieh B W, C.C.H., Chang S C., Scheduling semiconductor wafer fabrication by using ordinal optimization-based simulation[J]. IEEE Transactions on Robotics and Automation, 2001. 17(5), P599-608.
- [4] Shen Y X, L.R.C., Stochastic wafer fabrication scheduling[J]. IEEE Transactions on Semiconductor Manufacturing, 2003. 16(1), P 2-14.
- [5] Chiang T C, H.A.C., Fu L C, Modeling, scheduling, and performance evaluation for wafer fabrication: a queueing colored Petri-net and GA-based approach[J]. IEEE Transactions on Automation Science and Engineering, 2006.3(3), P330-338.
- [6] Bertsekas, D.P.a.J.N.T., Neuro-dynamic programming[M]. 1996: Athena scientific.
- [7] Jin huiyu, Application of Neuro-dynamic programming in Re-entrant Lines[D], Shenyang Automatic Institute.2001.(Ch) 金辉宇, 神经元动态规划在可重入生产系统调度中的应用[D], 中国科学院沈阳自动化研究所. 2001.
- [8] Wang Ying, Zhu Shunzhi, Xu Wei, Simulation Framework of Scheduling Re-entrant lines with Neuro-Dynamic Programming[J]. Information and Control, 2007. 36(2), P218-224.(Ch). 王颖, 朱顺志, 许威, 基于神经元动态规划的可重入生产系统调度的仿真框架[J]. 信息与控制, 2007. 36(2), P218-224.
- [9] Puterman, M.L., Markov decision process discrete stochastic dynamic programming[M]. 2005: New Jersey: John Wiley&Sons.
- [10] Wang Ying, Z.S., Li MQ, Scheduling of Re-entrant Lines with Neuro-Dynamic Programming Based on a New Evaluating Criterion[J]. Lecture Notes In Computer Science, 2006. 3973,P 921-926.
- [11] Baoqun Yin, Queueing system performance analysis and Markov control processes[M], Press of USTC, Hefei, 2003.(Ch) 殷保群, 排队系统性能分析与 Markov 控制过程[M]. 2003: 中国科学技术大学出版社.
- [12] Kumar, S.a.P.R.K., Performance bounds for queueing networks and scheduling policies[J]. IEEE Transactions on Automatic Control, 1994. 39(8), P1600-1611.
- [13] Wein, L.M., Scheduling semiconductor wafer fabrication[J]. IEEE Transactions on Semiconductor Manufacturing, 1988. 1(3), P115-129.
- [14] Cai Qiaoru, A Real-Time Scheduling and Rescheduling System Based on RFID for Semiconductor Foundry Fabs[J]. Journal of Industry and engineering, 2007. 24(6), P437-445.
- [15] Han-Pang Huang, C.-T.W. Modeling and Performance Evaluation for Automated Material Handling Systems in a 300mm Foundry Fab[C]. International Conference on Robotics & Automation. 2002.20(2), P91-100.
- [16] Li li, Jiang Hua, Wu Qidi, Research on dynamic dispatching rule for semiconductor wafer fabrication[J]. Computer Integrated Manufacturing Systems,2004. 10(8), P945-954.(Ch). 李莉, 乔., 姜桦, 吴启迪, 半导体生产线动态调度方法研究. 计算机集成制造系统[J], 2004. 10(8), P945-954.
- [17] Li li, Dynamic real time dispatching rule for semiconductor wafer fabrication[D]. Tongji University. 2005.(Ch) 李莉, 半导体生产线动态实时智能调度方法研究[D] 同济大学. 2005.