

基于模糊神经网络的半导体生产线 重调度策略优化*

张洁 翟文彬 严隽琪 马登哲

(上海交通大学计算机集成制造研究所 上海 200030)

摘要：针对缺乏半导体生产线重调度策略优化方法的研究现状，提出了基于模糊神经网络的半导体生产线重调度策略优化技术。将重调度策略划分为半导体生产线、设备组和设备重调度层次，利用仿真评价确定优化的重调度策略，并获得样本数据。通过对模糊神经网络训练，建立干扰和半导体生产线状态等输入参数与优化的重调度策略输出之间的映射关系。以上海某 125 mm 晶圆生产线为例，结果表明了该重调度策略优化方法的有效性。

关键词：半导体生产线 重调度 模糊神经网络

中图分类号：TH166

0 前言

半导体生产线属于可重入生产系统，具有可重入流、单件与批处理设备组共存、设备负载不均衡等不同于传统生产线的特殊性和复杂性^[1]。半导体加工设备随机故障、晶圆工件重加工和紧急定单等不确定因素，使得半导体生产线处于高度不确定性环境中。为了提高系统的生产效率和稳定性，必须根据干扰对系统的影响，进行有效的重调度。

重调度研究主要包括启发式方法和智能方法。H. Toba^[2]提出了基于分段法的实时反应的重调度启发式算法，用于半导体生产线的重调度。R. J. Abumaizar 等^[3]提出了随机扰动下的 Jobshop 重调度启发式算法，通过对受干扰的工序的调整，实现对调度的修正。针对 Jobshop 重调度问题，H. Li^[4]集成仿真、人工神经网络、专家系统和派工规则，提出了 Jobshop 生产重调度专家仿真系统。

综上所述，重调度问题研究对象集中在 Jobshop 生产系统，而可重入生产系统重调度技术局限于单一启发式算法，目前缺乏半导体生产线重调度策略优化技术的研究。针对以上不足，侧重于基于模糊神经网络 (Fuzzy neural networks, FNN) 的半导体生产线重调度策略优化方法的研究。

1 半导体生产线重调度问题分析

1.1 干扰分析

半导体生产线处于不确定的动态调度环境，其

运作过程中的干扰主要包括设备故障、工件重加工和紧急定单。当干扰发生时，需要根据当前的系统运行状态和干扰对调度的影响，选择优化的重调度策略，以确保半导体生产线调度的稳定性和有效性。为了便于定量地描述干扰对半导体生产线调度的影响，可以把它们统一转化为对设备加工工时的影响^[3]。因此，干扰对调度的影响可描述为以下几个方面。

设备故障干扰：假设 M_{1j} 表示发生故障的设备组 j ； m_{ji} 为设备组 j 中，发生故障的设备 i ； t_{1ji} 为设备组 j 中，设备 i 的故障修理时间，单位为 h ； t_1 为由于设备故障导致的系统加工能力减少的工时，即 $t_1 = \sum_{M_{1j} \in F_1} \sum_{m_{ji} \in M_{1j}} t_{1ji}$ ，其中， F_1 为发生故障的设备组集合。

工件重加工干扰：假设 M_{2j} 表示加工重加工工件的设备组 j ； R_{2j} 表示设备组 j 重加工工件集合； p_{jk} 表示设备组 j 中需要重加工的工件 k ； t_{2ik} 为工件 k 的重加工时间，单位为 h ； t_2 为由于工件重加工导致的系统加工能力增加的需求，即 $t_2 = \sum_{M_{2j} \in F_2} \sum_{p_{ik} \in M_{2j}} t_{2ik}$ ，

其中， F_2 为发生工件重加工的设备组集合。

紧急定单干扰：假设 M_{3j} 表示当前调度计划区内，紧急定单涉及的设备组 j ； R_{3j} 表示紧急定单涉及的设备组 j 加工的工件集合； p_{jk} 表示紧急定单中设备组 j 加工的工件 k ； t_{3jk} 为紧急定单中工件 k 的加工时间，单位为 h ； t_3 为由于紧急定单导致的系统加工能力增加的需求，即 $t_3 = \sum_{M_{3j} \in F_3} \sum_{p_{ik} \in M_{3j}} t_{3ik}$ ，

其中， F_3 为紧急定单相关的设备组集合。

干扰对半导体生产线调度的影响为

* 国家自然科学基金资助项目 (50375098) 20041112 收到初稿，20050329 收到修改稿

$$t = t_1 + t_2 + t_3$$

1.2 半导体生产线重调度策略

分析半导体生产线的实时状态和意外事件干扰对调度的影响,进行分层次重调度(图1)。当干扰对调度影响较大时,调度需要进行全局协调。此时,采用半导体生产线全局重调度策略,根据系统当前状态,调用半导体生产线短期重调度算法^[5],调整全局调度。然后,根据全局调度生成的设备组调度目标,设备组调用半导体生产线动态调度算法,实现设备的实时调度;当干扰对调度影响较小时,采用设备重调度策略,此时,调度计划的晶圆加工顺序不变,开始加工时间顺延干扰占用的工时;当干扰对调度影响介于以上两者之间时,采用设备组重调度策略,此时,根据全局调度生成的设备组日调度目标不变,受干扰的设备组调用半导体生产线动态调度算法,实现调度局部调整。

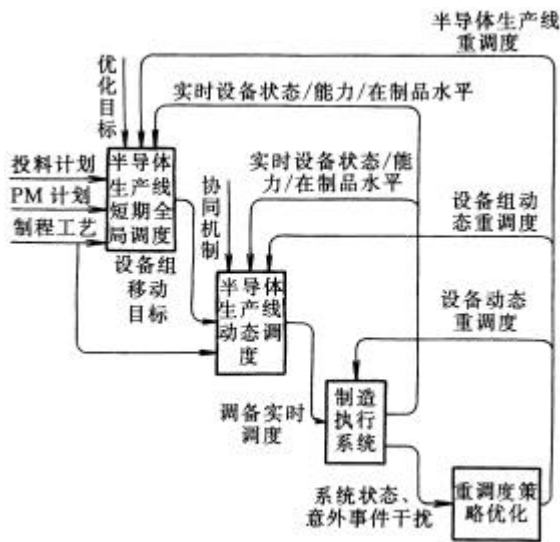


图1 半导体生产线重调度策略

1.3 重调度策略评价

重调度策略的目标是尽可能减小干扰对调度的扰动,同时使调整后的调度依然具有较好的性能。本文采用有效性和稳定性评价重调度策略的健壮性。

有效性表示调整前、后半导体生产线调度的目标函数值变化量比较原始调度目标函数值的相对变化率,即 $a = \frac{O_n - O_0}{O_0}$, 其中, O_n 为调整后调度的目标函数值, O_0 为调整前调度的目标函数值。

稳定性表示调整前、后半导体生产线调度的晶圆开始加工时间的平均变化率,即 $b = \frac{\sum_{(i,s) \in S} |t'_{is} - t_{is}|}{\sum_{(i,s) \in S} q_{is} t_p}$,

其中, t'_{is} 为调整后产品 i 工段 s 的晶圆开始加工时间, t_{is} 为调整前产品 i 工段 s 的晶圆开始加工时间, S 为全局调度生成的半导体生产线调度集, q_{is} 为产品 i 工段 s 的晶圆数量, t_p 为调度时区长度。

重调度策略的健壮性为 $R = wa - (1-w)b$, 其中, $0 < w < 1$ 为权重系数。

2 基于 FNN 的重调度策略优化

2.1 FNN 的输入层变量

FNN 的输入层变量由半导体生产线状态参数和干扰参数 t 组成。其中, 半导体生产线状态参数包括以下几个方面。

受干扰的设备组平均队列长度 L 反映了设备组 j 的空闲状态, 即, $L = \sum_{M_j \in (F_1 \cup F_2 \cup F_3)} L_j / R$, 其中, L_j

为设备组 M_j 的队列长度, R 为受干扰的设备组数。

当前半导体生产线调度稳定性 b_c 为开始加工或半导体生产线全局重调度到现在, 由全局调度生成的调度目标与实际调度结果之间的平均开始加工时间偏差值, 即 $b_c = \sum_{\substack{(i,s) \in S \\ t_{cis} \leq t_c}} |t'_{cis} - t_{cis}| / \sum_{\substack{(i,s) \in S \\ t_{cis} \leq t_c}} q_{is}$, 其中, t'_{cis}

为产品 i 工段 s 晶圆的实际开始加工时间, t_{cis} 为全局调度生成的产品 i 工段 s 晶圆调度计划的开始加工时间(取为所处时段的中点时刻), t_c 为当前干扰发生时刻。

受干扰的设备组平均相对负载 h 为当前干扰发生时刻 t_c 到计划终止时刻 t_e 期间, 不考虑干扰工时的设备组平均相对负载, 即 $h = \sum_{\substack{(i,s) \in S_d \\ t_e \geq t_{is} \geq t_c}} t_{pis} / \sum_{M_j \in (F_1 \cup F_2 \cup F_3)} n_j (t_e - t_c)$,

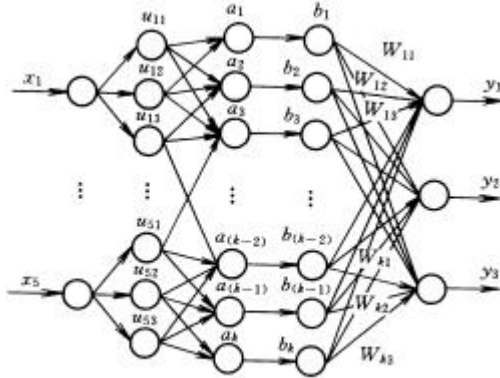
其中, t_{pis} 为产品 i 工段 s 晶圆的加工工时, t_{is} 为产品 i 工段 s 晶圆的开始加工时间, S_d 为全局调度生成的调度中受干扰的设备组的调度集, n_j 为设备组 M_j 的设备数量。

受干扰的设备组工件后序平均松弛时间 t_s 表示干扰发生时, 考虑后序工件加工情况下, 设备组调度的可调整程度, 即 $t_s = \sum_{\substack{(i,s) \in S_d \\ t_e \geq t_{is} \geq t_c}} (t_{i(s+1)} - t_{is} - t_{pis}) / \sum_{\substack{(i,s) \in S_d \\ t_e \geq t_{is} \geq t_c}} q_{is}$ 。

2.2 FNN 的结构

FNN 包括 5 层, 采用 Mamdani 推理法。设描述输入输出关系的模糊规则为: R_i : 如果 x_1 是 A_{1i} 和 x_2 是 A_{2i} 和... 和 x_m 是 A_{mi} 那么 y_1 是 B_{1i} 和 y_2 是 B_{2i} 和... 和 y_m 是 B_{ki} , $i=1,2,\dots,n$, 其中, n 为规则总数, m 为输入变量数, k 为输出变量数, A_{ji} 为 x_j 的模糊语言变量值, B_{ji} 为 y_j 的模糊语言变量值。

对 FNN 的具体描述如下所述(图2)。



第 1 层 第 2 层 第 3 层 第 4 层 第 5 层

图 2 FNN 结构

第 1 层为输入层，输入矢量为 $X=[x_1, x_2, x_3, x_4, x_5]=[w, b_c, h, t_s, t]$ ，结点输入输出函数为 $f_{i(1)} = x_{i(0)} = x_i; x_{i(1)} = g_{i(1)} = f_{i(1)}; i=1,2,L,5$ 。

第 2 层为模糊化层，用于计算各输入分量属于各语言变量值模糊集合的隶属度。隶属度函数采用高斯函数表示的铃形函数，即 $u_{ij} = \exp\left[-\frac{(x_i - c_{ij})^2}{s_{ij}^2}\right]$ ， $i=1,2,L,5, j=1,2,L, k_i$ ， c_{ij} 和 s_{ij} 分别为隶属度函数的中心和宽度变量。结点输入输出函数为

$$f_{ij(2)} = -\frac{(x_i - c_{ij})^2}{s_{ij}^2}$$

$$x_{ij(2)} = u_{ij} = g_{ij(2)} = \exp f_{ij(2)} = \exp\left[-\frac{(x_i - c_{ij})^2}{s_{ij}^2}\right]$$

$$i=1,2,L,5 \quad j=1,2,L, k_i$$

第 3 层为规则层，每个结点代表一个模糊规则，用于匹配模糊规则的前件，计算出每条规则的适用度，即 $a_j = \prod_{l=1}^5 u_{li}(x_{li})$ ， $j=1, 2,L, n$ 。结点输入输出函数为

$$f_{j(3)} = \prod_{l=1}^5 x_{li(2)} = \prod_{l=1}^5 u_{li}(x_{li})$$

$$x_{j(3)} = g_{j(3)} = f_{j(3)} \quad j=1,2,L, n$$

第 4 层为量纲一化层，结点数与第 3 层相同，用于规则适用度的量纲一化计算，即 $b_j = a_j / \sum_{i=1}^n a_i, j=1, 2,L, n$ 。结点输入输出函数为 $f_{j(4)} = x_{j(3)} / \sum_{i=1}^n x_{i(3)} = a_j x_{j(4)} / \sum_{i=1}^n a_i = g_{j(4)} = f_{j(4)}, j=1,2,L, n$ 。

第 5 层输出层，实现输出变量的去模糊化计算，结点数等于重调度策略的数量。当选用某个调度规则时，与它对应的输出接近 1，否则接近 0。结点输入输出函数为

$$f_{i(5)} = \sum_{j=1}^n w_{ij} x_{j(4)} = \sum_{j=1}^n w_{ij} b_j$$

$$x_{j(5)} = g_{j(5)} = f_{j(5)} \quad i=1,2,3$$

式中 w_{ij} ——连接权参数

2.3 学习算法

FNN 的学习包括结构学习和参数学习。首先进行结构学习，根据半导体生产线干扰的统计分布(表 1)，建立半导体生产线仿真模型。仿真过程中，当干扰发生时，基于有效性和稳定性指标评价重调度策略的健壮性，选择健壮性最好的重调度策略，建立半导体生产线状态参数、干扰大小与重调度策略之间的对应关系，实现 FNN 样本建立。按照重调度策略将样本分类，采用模糊 c-均值聚类法对每类输入样本空间进行聚类划分，若训练样本被分为 n_i 类，则对应 n_i 个模糊规则，总规则数为 $n = \sum n_i$ 。聚类的中心和半径确定了 FNN 输入变量隶属度函数参数 c_{ij} 和 s_{ij} 的初值，从而实现 FNN 的模糊输入空间划分和规则提取。

设样本的输出误差代价函数为 $E = \frac{1}{2} \sum_{i=1}^3 (y'_i - y_i)^2$ ，

其中 y'_i 为期望输出， y_i 为实际输出。参数学习采用误差反传算法，利用有监督的梯度下降寻优算法来调节 FNN 的参数 w_{ij} 、 c_{ij} 、 s_{ij} ，最后得到的参数调整公式为

$$w_{ij}(k+1) = w_{ij}(k) - g(y'_i - y_i) b_j$$

$$c_{ij}(k+1) = c_{ij}(k) - g d_{ij} \frac{2(x_i - c_{ij})}{s_{ij}^2}$$

$$s_{ij}(k+1) = s_{ij}(k) - g d_{ij} \frac{2(x_i - c_{ij})^2}{s_{ij}^3}$$

式中 g ——学习率，且 $g > 0$

$$d_{ij} = \sum_{k=1}^n \left[\sum_{g=1}^3 (y'_k - y_k) w_{gk} \frac{\sum_{i=1}^n a_i}{\sum_{i=1, i \neq k}^n a_i} \right] s_{ij} \times \exp\left[-\frac{(x_i - c_{ij})^2}{s_{ij}^2}\right]$$

$$s_{ij} = \begin{cases} \prod_{\substack{i=1 \\ i \neq j}}^5 u_{li}(x_{li}) & g_{ij}^{(2)} = u_{ij} \\ 0 & \text{其他} \end{cases}$$

u_{ij} ——第 j 个规则结点的一个输入

3 实例分析

采用上海某半导体制造企业 6 寸晶圆生产线的

实际生产数据,利用面向对象仿真软件 Emplant,建立仿真模型。该半导体生产线由 11 个经常成为瓶颈的设备组构成,共 34 台关键设备,3 种晶圆产品

共 85 个工段。通过仿真获得采用 3 种重调度策略的各 50 组数据,其中每种分别随机抽取 30 组数据作为训练样本,其余 20 组数据作为测试数据(表 1,2)。

表 1 训练样本

样本号	受干扰设备组 平均队列 x_1	调度稳定性 x_2/h	受干扰设备组 平均负荷 x_3	受干扰设备组后序 平均松弛时间 x_4/h	干扰 x_5/h	重调度策略		
						设备层 y_1	设备组层 y_2	生产线层 y_3
1	1	1.10	0.59	6.42	2.14	1	0	0
2	2	0.78	0.52	6.51	2.01	1	0	0
3	0	0.81	0.46	5.16	1.76	1	0	0
M	M	M	M	M	M	M	M	M
41	3	3.22	0.73	6.32	7.32	0	1	0
42	2	2.24	0.74	4.84	6.04	0	1	0
43	3	4.19	0.64	2.11	6.46	0	1	0
M	M	M	M	M	M	M	M	M
88	6	4.23	0.75	3.81	9.81	0	0	1
89	4	4.15	0.76	5.64	9.18	0	0	1
90	5	4.01	0.76	4.97	9.77	0	0	1

表 2 测试结果

样本号	受干扰设备 组平均队列 x_1	调度稳 定性 x_2/h	受干扰设备组 平均负荷 x_3	受干扰设备组 后序平均松弛 时间 x_4/h	干扰 x_5/h	重调度策略目标			FNN 输出的重调度策略		
						设备层	设备组层	生产线层	设备层	设备组层	生产线层
						y_1	y_2	y_3	y_1	y_2	y_3
1	0	1.07	0.48	5.81	1.21	1	0	0	0.974 9	-0.037 5	0.072 7
2	2	1.42	0.49	4.62	1.42	1	0	0	1.074 4	-0.190 5	0.119 2
3	1	1.29	0.52	5.54	1.79	1	0	0	1.005 1	-0.002 1	0.001 3
M	M	M	M	M	M	M	M	M	M	M	M
15	3	3.18	0.76	6.13	6.42	0	1	0	0.026 1	0.894 4	0.090 2
16	3	3.34	0.76	5.73	5.76	0	1	0	0.022 2	0.955 7	0.028 3
17	3	2.21	0.81	4.53	6.11	0	1	0	-0.025 3	1.070 0	-0.037 6
M	M	M	M	M	M	M	M	M	M	M	M
58	8	4.67	0.77	3.21	8.94	0	0	1	0.433 9	-0.025 3	0.621 7
59	6	3.96	0.84	5.18	9.46	0	0	1	0.017 9	0.266 2	0.776 6
60	7	3.75	0.72	4.76	9.71	0	0	1	-0.071 5	-0.009 0	1.145 7

基于 FNN 的半导体生产线重调度策略优化算法采用 Matlab6.5 编程实现。FNN 中,采用模糊 C—均值聚类法获得了 12 个规则。FNN 训练时,学习率 g 取为 0.01, w_{ij} 初值均取为 1,网络训练过程如图 3

所示。通过 FNN 输出与相应目标的线性回归分析,验证了重调度策略优化算法的有效性。

4 结 论

半导体加工设备随机故障、晶圆工件重加工和紧急订单等不确定因素,使得半导体生产线处于高度不确定性环境中。为了提高系统的生产效率和稳定性,必须根据干扰对系统的影响进行有效的重调度。针对缺乏半导体生产线重调度策略优化方法的研究现状,提出了基于模糊神经网络的半导体生产线重调度策略优化技术。以上海某半导体制造企业 6 英寸晶圆生产线为实例,通过 FNN 输出与相应目标的线性回归分析,验证了重调度策略优化算法的有效性。

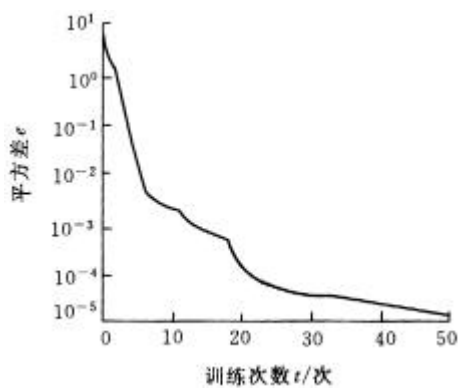


图 3 FNN 学习曲线

参 考 文 献

- 1 Uzsoy R, Lee C Y, Martin-Vega L A. A review of production planning and scheduling models in the semiconductor industry part II : shop-floor control. IIE Transactions, 1994, 26(5) : 44 ~ 55
- 2 Toba H. Segment-based approach for real-time reactive rescheduling for automatic manufacturing control. IEEE Transactions on Semiconductor Manufacturing, 2000, 13(3) : 264 ~ 272
- 3 Abumaizar R J. Rescheduling job shops under random disruptions. International Journal of Production Research, 1997, 35(7) : 2 065 ~ 2 082
- 4 Li H. Production rescheduling expert simulation system. European Journal of Operational Research, 2000, 124(2) : 283 ~ 293
- 5 翟文彬, 褚学宁, 张洁, 等. 基于组合拍卖的半导体生产线短期调度技术研究. 机械工程学报, 2004, 40(9) : 95 ~ 99

FUZZY NEURAL NETWORKS BASED RESCHEDULING STRATEGIES OPTIMIZATION OF SEMICONDUCTOR FABRICATION LINE

Zhang Jie Zhai Wenbin Yan Juanqi

Ma Dengzhe

(Institute of Computer Integrated Manufacturing,
Shanghai Jiaotong University, Shanghai 200030)

Abstract: For the lack of rescheduling optimization approach for semiconductor fabrication line status in quo, a novel fuzzy neural networks based algorithm for rescheduling strategies optimization of semiconductor fabrication line is proposed. The rescheduling strategies are divided into line layer, machine group layer and machine layer. The samples for FNN are setup by simulation evaluation. The relation between the input of FNN, such as disturbance, system state parameters, and output of FNN, optimal rescheduling strategy, is built by FNN. With an illustrative example of a 152 mm semiconductor fabrication line in shanghai, the lineal regression between FNN output and targets demonstrates the effectiveness of the approach.

Key words : Semiconductor fabrication line Rescheduling

Fuzzy neural networks

作者简介: 张洁, 女, 1963 年出生, 教授. 主要从事制造系统建模与仿真、调度与控制等研究。

E-mail : zwbin_2000@163.net