Vol. 43 No. 10 Oct. 2015

半导体生产线基于 DBR 和 ANFIS 相融合的 动态调度方法研究

曹政才^{1,2},彭亚珍^{1,2},李 博³,刘 民⁴

(1.北京化工大学信息科学与技术学院,北京 100029;2.吉林大学符号计算与知识工程教育部重点实验室,吉林 130012; 3.中国科学院理化技术研究所,北京 100190;4.清华大学自动化系,北京 100084)

摘要:实际半导体制造过程调度问题具有大规模、不确定、带复杂约束及多目标等综合复杂性,要确保上述生产过程优化运行,必须及时有效进行动态调度.鉴于半导体生产线具有多重人特征及紧急订单对常规订单产生影响,本文给出一种基于"鼓-缓冲-绳子"(Drum-Buffer-Rope, DBR)和自适应模糊推理系统(Adaptive Neuro-Fuzzy Inference System, ANFIS)相融合的半导体生产线动态调度方法.首先以最大化瓶颈设备有效产能及保证生产线负荷均衡为目标,将投料控制与工件调度有机结合进行 DBR 优化算法设计;其次根据生产线运行过程中积累大量的历史数据与实时数据,利用 ANFIS 构建紧急订单相关信息预测模型;再次结合专家经验知识,利用模糊推理系统将预测结果与相应的DBR 算法相融合,使生产线提前调整投料策略,保证紧急订单到来时生产线能够有效完成其加工任务及减小紧急订单与常规订单之间的相互影响;最后通过某半导体生产线进行仿真验证,该方法能够实现生产线的多目标优化,为解决实际半导体生产调度问题提供参考.

关键词: 半导体生产线; 动态调度; 紧急订单预测; 鼓-缓冲-绳子

中图分类号: TP273 文献标识码: A 文章编号: 0372-2112 (2015)10-2082-06 电子学报 URL: http://www.ejournal.org.cn DOI: 10.3969/j.issn.0372-2112.2015.10.030

A DBR and ANFIS Integrated Dynamic Scheduling Algorithm for Semiconductor Wafer Fabrication

CAO Zheng-cai^{1,2}, PENG Ya-zhen^{1,2}, LI Bo³, LIU Min⁴

- (1. College of Information Science and Technology, Beijing University of Chemical Technology, Beijing, China, 100029;
- 2. Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, Jilin University, Changchun, China, 130012;
 - 3. Technical Institute of Physics and Chemistry of the Chinese Academy of Sciences, Beijing, China, 100190;
 - 4. Department of Automation , Tsinghua University , Beijing , China , 100084)

Abstract: The scheduling problem of actual semiconductor production process is of large scale, uncertainty, complicated constraints and multiple objectives, and the optimization of the process depends upon a timely and efficient dynamic scheduling. Considering the multi-reentrant characteristic of semiconductor wafer fabrication and the influence of rush orders upon regular orders, a Drum-Buffer-Rope (DBR) theory and Adaptive Neuro-Fuzzy Inference System (ANFIS) integrated dynamic scheduling algorithm is proposed. Firstly, a DBR optimization algorithm based on the combination of releasing control and lots scheduling is designed, aiming at maximizing the effective capacity of bottleneck equipment and ensuring the wafer fabrication load in a balance. Secondly, using the historical and real-time data obtained in the wafer fabrication, a prediction model of rush orders is constructed on the basis of ANFIS. Thirdly, the prediction result is integrated with the DBR algorithm combined with the experts' knowledge and practice experiment, to adjust the wafer fabrication before rush orders come, which guarantees that the rush order could be completed effectively, and reduces the interplay between the rush order and the regular one. The proposed scheduling algorithm is simulated on a certain semiconductor wafer fabrication system, and the result shows that the method can realize the multi-objective optimization and provide reference to the practical semiconductor wafer fabrication scheduling.

Key words: semiconductor wafer fabrication; dynamic scheduling; rush order prediction; drum-buffer-rope

1 引言

实际半导体制造过程具有大规模(机器数达上百 至上千、工件数达数百至数千,操作数达数千至上万)、 不确定(工件的释放时间、加工时间和交货期及加工能 力等)、带复杂约束(批加工、资源约束及工艺约束等) 及多目标等综合复杂性,要确保上述生产过程优化运 行,必须及时有效进行动态调度,该方向的研究近年来 一直受到学术界和工业界众多学者的极大关注[1,2]. 半 导体生产线涉及工件、设备数量多及多重入特性,研究 常用启发式规则、计算智能、数据挖掘和离散事件仿真 等方法求解,如:借助干组合规则获得性能较优的调度 方案[3,4]:文献[5]根据虚拟生产线,在制品数和生产时 间被分配到产品的每个加工步骤上,给出一种拉式的 投料策略和派工规则控制每一条虚拟生产线,解决半 导体制造系统多目标调度问题: Oiao 等[6]采用分层有色 Petri 网来描述半导体生产线加工状态、性能及路径约 束,采用改进的遗传算法快速搜索规则组合,求得调度 问题的近似最优解; Li 等[7] 基于生产过程调度相关历 史数据和生产过程仿真模型提取少量数据对,并采用 整体模糊化和数据趋势预测方法生成大量用于获取启 发式调度规则选择策略的训练数据对,应用 BP 神经网 络进行调度规则的挖掘; Meidan 等[8]提出一种基于机器 学习和数据挖掘相结合的方法来识别和预测半导体生 产线关键因素对加工周期的影响,该方法主要采用条 件最大化信息共享方法来识别影响加工周期最大的因 素,然后采用贝叶斯分类器进一步选择最小的、最有影 响的关键因素集合来进行工件加工时间预测.

迄今生产过程调度方法研究上已取得很多理论成果,但仍难以胜任实际调度问题的求解.DBR方法将瓶颈作为制造系统的鼓,以其节奏带动生产,用缓冲器充分保证瓶颈设备的利用率;通过绳子驱动系统其它非瓶颈设备,使其与瓶颈设备同步^[9].文献[10]基于系统瓶颈的识别,把整个生产线划分成若干条"层生产线",并将各层起始处的系统瓶颈看成投料点,根据层瓶颈对其投料进行控制;KOH等^[11]对比生产线有效产出率

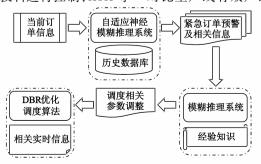


图1 半导体生产线动态调度实施框图

及最优净利润等指标,采用 DBR 思想的投料策略优于固定在制品投料;文献[12]采用 DBR 方法,着重考虑瓶颈调度,通过安排工件在瓶颈上的各工序加工时间,指导实际生产.

紧急订单是半导体制造企业中常见的一类订单,在加工过程中具有较高的加工优先级,且容易对生产线稳态产生负面影响.目前,针对紧急订单的研究还不多见,Chen等[13]提出基于启发式规则的紧急订单模型;Wang等[14]构建一个供应链网络优化模型,以合适的方式将紧急订单分发到供应链网络,以满足客户对紧急订单的要求.

借鉴 DBR 理论,将投料控制与工件调度有机结合进行优化调度算法设计;采用 ANFIS 技术对紧急订单进行预测及结合专家经验知识,来确定基于 DBR 的调度算法中的相关参数,以此实现对调度过程的提前调整,保证紧急订单顺利完成,来提高半导体生产线运行效率.图1.半导体生产线动态调度实施框图.

2 基于扩展 DBR 的生产线优化调度算法

DBR 方法的三要素分别是^[10]:1.鼓(Drum):瓶颈资源;2.缓冲器(Buffer):借助缓冲管理最大化利用瓶颈资源;3.绳子(Rope):利用瓶颈资源的节奏控制非瓶颈资源的调度.

2.1 瓶颈识别[2]

瓶颈即生产系统的约束,能最大程度影响到系统的生产速率.

(1)主瓶颈识别

制造系统的核心约束被认为是系统的主瓶颈.在 调度期订单确定的基础上,依据工件的工艺计算,可以 得到各加工中心的相对负荷,由此可以确定系统瓶颈, 具体计算公式为:

$$\max b_m = \max \left\{ \frac{D_m}{C_m} \right\} = \max \left\{ \frac{1}{C_m} \sum_{n=1}^N \omega_{nm} Q_n \right\}, \quad m = 1, 2, \dots M$$
(1)

式中: D_m 为计划期内资源 m 的累计负荷; ω_{nm} 为产品 n 关于资源 m 的消耗系数; Q_n 为产品 n 的需求数量; C_m 为计划期内资源 m 的最大生产能力; b_m 为资源 m 的负荷率; M 为生产系统中的资源总数; N 为产品种类.

(2)次瓶颈识别

设备故障等突发事件可能会引起局部负载不均衡,进而导致生产线性能降低甚至引起不必要的瓶颈漂移.为降低该问题造成的影响,给出次瓶颈的概念.利用各设备前缓冲区长度识别次瓶颈,由于生产线的动态特性,具体采用如下方法来识别:

①引入指数变换并定义 $IBD_m(t)$ 为加工中心 m 的 瞬时瓶颈度:

$$IBD_m(t) = E(QL_m(t)) = 1 - e^{-\lambda \cdot QL_m(t)/C_m}$$
 (2)

式中: λ 为与生产线特性相关的常量,通过设置该常量可以将瓶颈度转换到一个合适的范围; $QL_m(t)$ 为加工中心 S 在 t 时刻的缓冲长度,缓冲长度指对应缓冲区内工件的剩余平均加工时间.

②为兼顾数据的静态性能和动态性能,进一步采用指数平滑方法对瞬时瓶颈度平滑化得到 t 时刻加工中心m 的瓶颈度:

$$BD_{m}(t) = \alpha \cdot IBD_{m}(t) + (1 - \alpha) \cdot BD_{m}(t - T)$$
(3)

其中: $BD_m(t)$ 为对应 t 时刻瓶加工中心 m 的颈度; T 为 更新瓶颈度的时间间隔; α 为平滑系数, 其设置与生产工况波动程度及更新时间间隔 T 有关. 将 $BD_m(0)$ 初始 化为 0, 而后 $BD_m(t)$ 可以通过(3)式计算得到.

③当且仅当满足下述条件时:

$$BD_{mb}(t) = \max_{m} BD_{m}(t) \tag{4}$$

设备 mb 为系统的瓶颈.

2.2 瓶颈的充分利用

鉴于半导体制造系统的多重入特性,需要从两方面予以考虑:避免瓶颈设备饥饿及保证各瓶颈工序相对占用率一致.

(1)避免瓶颈饥饿

如图 2 所示为生产线层划分示意图,鉴于缓冲区内工件情况难以预测,通过保证层负荷处于一定水平来防止瓶颈饥饿.每一层定义为从投料工序开始或者瓶颈工序结束起到下一道瓶颈工序结束或生产最后一道工序结束之间工序的集合.

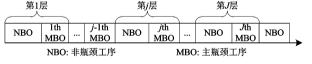


图 2 生产线层划分示意图

(2)最大化有效产出

并非瓶颈加工都会转化为有效产出,因此在调度过程中需要充分考虑到主瓶颈上个工序的占用比例,即考虑各瓶颈工序间的负荷平衡.只有各瓶颈工序间达到相对平衡状态,才能使有效产出尽可能的接近瓶颈设备的理想产能.

2.3 配合瓶颈的节奏

在 DBR 调度方法中, 鼓是瓶颈的节奏, 整个生产线按照鼓的节奏生产. 为防止不确定性因素对瓶颈造成影响, 对于每一层生产线设定层时间缓冲, 以保护对应层内的瓶颈工序. 通过次瓶颈管理, 提升其后道工序优先级及降低其前道工序优先级, 防止该局部干扰进一步影响到其他加工中心而造成整个生产线的波动.

2.4 调度算法

根据 DBR 理论,在获得半导体制造系统瓶颈后,需要以瓶颈为控制核心,充分利用瓶颈资源,以此带动其它设备的调度,实现制造系统有序的生产,具体调度算法如下^[2,9]:

设 $i(i=1,\cdots,I)$ 为工件的第 i 道工序, $BN_j(j=1,\cdots,J)$ 为工件的第 j 道瓶颈工序, SN_k 为工件的第 k 道次瓶颈工序.工件的复合优先级 P_i 由以下 5 个子优先级组成:

(1)瓶颈优先级 P_{Mi} :该优先级为瓶颈工序优先级, 其主要与上下游层负荷及当前工序缓冲长度有关,可 计算如下:

$$P_{Mi} = \begin{cases} m_i (L_j / L_0 - 1) + m_0 B_i + m_2 (L_{i+1} / L_0 - 1), \\ i = B N_j \text{ and } 1 \leq j < J \\ m_1 (L_J / L_0 - 1) + m_0 B_i, i = B N_J \\ 0, & \text{else} \end{cases}$$
 (5)

式中: m_0 , m_1 , m_2 均为经验常量; B_i 为工序 i 的缓冲长度,对应工序为 i 的待加工工件剩余平均加工时间; L_0 为期望层负荷.

(2)非瓶颈加工优先级 P_N :非瓶颈工序的加工应与瓶颈工序相配合,该优先级主要与当前工序缓冲区长度、当前层负荷及对应层内瓶颈工序缓冲长度有关:

$$P_{Ni} = \begin{cases} n_0 B_i + n_1 (L_j / L_0 - 1) + n_2 B_{BN_j}, \\ & i \in jth \ layer \ and \ i \neq BN_j \\ n_0 B_i & i > BN_J \\ 0, & \text{else} \end{cases}$$
 (6)

对应的 n_0, n_1, n_2 为经验常数.

(3)次瓶颈优先级 P_{Si} ,为解决局部堵塞问题,引入次瓶颈优先级,以改变次瓶颈上下游工序优先级,防止局部问题向全局扩展.其计算方法如下:

$$P_{Si} = \begin{cases} s_1 B_{i+1}, & i+1 \in \{SN_k\} \\ s_2 B_{i-1}, & i-1 \in \{SN_k\} \\ 0, & \text{else} \end{cases}$$
 (7)

(4)工序优先级 P_{0i} : 采取依据鼓的节奏加工并采取由后至前的工序安排顺序设定优先级如下:

 $P_{0i} = p_1[(t - t_i)/D_t - 1] + p_2 \cdot i/I, i = 1, 2, \dots, I$ (8) 式中: t 为当前时间; t_i 为工序 i 的最近一次加工时间; p_1, p_2 为经验常数.

(5)投料优先级 P_{Ri} :利用瓶颈的负荷信息,转化为相应的优先级,该优先级主要影响到生产线是否投料的决策:

$$P_{Ri} = \begin{cases} r(1 - \sum_{j} \sum_{h} (\theta_{hj} \sum_{j_i = j}^{J} PT_{jl}) / BL), & i = 1 \\ 0 & \text{else} \end{cases}$$
(9)

式中:BL 为期望主瓶颈负荷;r 为权系数,将投料优先

级值调整到合适的范围.

最终,各工序复合优先级 P_i 可以通过如下公式计算得到:

$$P_i = P_{Mi} + P_{Ni} + P_{Si} + P_{Ri}, i = 1, 2, \dots, I$$
 (10)

通过使用该复合优先级,投料控制与工件调度被融为一起且考虑鼓的节奏,局部堵塞以及生产线信息.

3 基于 ANFIS 的紧急订单预测

ANFIS 是将模糊推理系统的先验知识概括、推理能力与神经网络的强大学习能力有机结合到一起,借助神经网络学习与信息存储能力,实现模糊推理系统自适应、自组织化^[15].

为实现半导体制造系统紧急订单预测,针对加工 n 类工件的生产过程构建如下紧急订单模型^[8]:

$$y = f(x_1, x_2, x_3, \cdots, x_{n+3})$$
 (11) 其中 x_1 为客户代码,其值为 $1 - m(m)$ 为整数,具有 m 位客户); $x_2, x_3, \cdots, x_{n+1}$ 表示各产品订单量,分别对应 n 种不同产品的需求数量; x_{n+2} 代表订单距交付期的时间; x_{n+3} 为紧急订单代码,其值为 1 表示该订单为紧急订单,0 表示该订单为常规订单; y 为 ANFIS 输出,输出为 0 表示下一订单为常规订单,否则下一订单是紧急订单,且其值代表了紧急订单的类型.同样在预测紧急订单量及紧急订单距交付期时间间隔时, y 分别指代紧急订单的量及紧急订单距交付期时间, 其具体流程如

- (1)载入数据: ANFIS 的参数需要通过训练数据进行反复的学习和更新;
- (2)采用减法聚类获取模糊规则,构造初始 FIS (Fuzzy Inference System,模糊推理系统)结构,隶属函数类型采用钟型函数;
- (3)确定训练选项值:采用混合学习算法优化模型 参数,设定训练周期和误差限;
 - (4)开始训练;

下[16]:

(5)待训练结束后观察训练数据或测试数据与模型输出之间的关系.

3 基于模糊推理系统的瓶颈负荷调整

模糊推理系统不需要通过机理分析来建立被控对象的精确数学模型,而是直接将人的先验知识转化成规则库,并通过模糊化与去模糊化将规则与实际应用相结合.通过模糊推理系统能有效地将紧急订单预测结果与先验知识相融合,使生产线瓶颈负荷状态在紧急订单到来之前提前得到调整,以降低紧急订单造成生产线的扰动作用.

通过分析生产调度系统可知紧急订单对准时交付 率及平均加工周期等参数影响重大,其中紧急订单量 及紧急订单交付期为紧急订单紧迫度的关键因素,在 DBR调度系统中,瓶颈负荷是对生产具有重要影响的 因素之一.为此设计模糊推理系统如下^[9]:

- (1)选用紧急订单量 OA 及其距交付期的时间 T 两个 ANFIS 预测量作为模糊推理系统的输入,将期望瓶颈负荷预留量 BL 作为输出,构造两输入单输出的模糊推理系统.
- (2)为充分考虑模糊推理系统的抗干扰能力采用 高斯隶属函数法模糊化、Mamdani 推理并采用重心法去 模糊化.
- (3)将输入输出论域分别划分为五个模糊子集 {NB,NS,Z,PS,PB},依据经验知识建立模糊推理规则 表,如表1所示:

表 1 模糊规则表

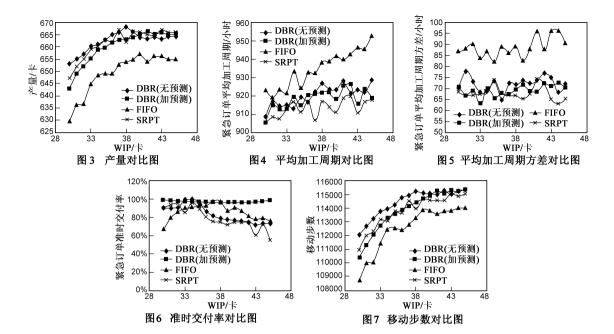
T	NB	NS	Z	PS	PB
NB	NB	NB	NB	NB	NB
NS	NS	Z	NS	NB	NB
Z	PS	PS	Z	NS	NB
PS	PB	PS	PS	Z	NS
PB	PB	PB	PB	PB	PB

5 仿真验证

以某半导体生产线模型为仿真对象进行验证.其中,仿真持续3年,每天24小时不间断加工,正式仿真前有半年时间进行预仿真.预仿真前90天投入无订单号的工件,投料固定时间间隔为36小时,使生产线达到饱和状态.第90至180天依据订单信息投料,使生产线进入稳定生产状态,投料固定时间间隔为38小时.整个预仿真期间工件调度采用FIFO策略,生产线故障维护采用基于定时维护的策略.

仿真结果如图 3 至 7 所示,图 3 为产量对比,从图中可以看出不加预测的 DBR 调度方法产量最高,而加入预测控制后,生产线为了调整相关参数状态,保证后续紧急订单的有效加工,因此会对常规订单造成一定影响,使得常规订单产量有所下降.

图 4 至图 6 为紧急订单的平均加工周期、平均加工周期方差及准时交付率,基于 DBR 的调度方法具有较好的性能,且加入紧急订单预测信息对生产线调节后,紧急订单性能指标都有所提升,尤其是准时交付率得到大幅度提高,且非常稳定.图 7 为移动步数,本文所给出调度方法在移动步数上有所降低,这也是保证紧急订单顺利加工的前提之一.仿真表明:在预测紧急订单相关信息的基础上提前调整生产线的瓶颈负荷期的方法能有效提升相应的生产线性能指标.



结论

- (1)给出扩展 DBR 优化调度方法,引入"层负荷"和 "次瓶颈"概念,通过局部与全局生产信息转化为相应 加工复合优先级,使非瓶颈配合瓶颈的节奏,实现瓶颈 产能最大化及生产线负荷均衡为目的.
- (2)构建基于 ANFIS 的半导体制造系统紧急订单预 测模型,在预测得到的紧急订单量及结合专家经验知 识,确定基于 DBR 的调度算法中的相关参数调整指标, 有效改善半导体生产线多性能指标.

参考文献

- [1] 刘民.基于数据的生产过程调度方法研究综述[J].自动 化学报,2009,35(6):785-806. Liu Min. A survey of data - based production scheduling methods[J]. Acta Automatica Sinica. 2009, 35(6): 785 - 806. (in Chinese)
- [2] Cao Z C, Peng Y Z, Wang Y J. A drum-buffer-rope based scheduling method for semiconductor manufacturing system [C]. IEEE Conference on Automation Science and Engineering, 2011:120 - 125.
- [3] Sarin S C, Varadarajan A, Wang L X. A survey of dispatching rules for operational control inwafer fabrication [J]. Production Planning and Control, 2011, 22(1):4-24.
- [4] Gloria R V, Eric T, Jacques P, et al. Dispatching of lots to dynamically reduce the wafers at risk insemiconductor manufacturing [C]. IEEE International Conference on Automation Science and Engineering, 2012, 920 – 923.
- [5] Li Y, Jiang Z B, Jia W Y. A pull VPLs based release policy and dispatching rule for semiconductor wafer fabrication [C].

- IEEE International Conference on Automation Science and Engineering, 2012:396-400.
- [6] Qiao F, Ma Y M, Li L, et al. A Petri Net and Extended Genetic Algorithm Combined Scheduling Method for Wafer Fabrication [J]. IEEE transaction on Automation Science and Engineering, 2013, 10(1): 197 - 204.
- [7] Li D C, Wu C S, Tasi T I, et al. Using mega-trend diffusion and artificial samples in small data set learning for early flexible manufacturing system scheduling knowledge [J]. Computers and Operations Research, 2007, 34(4): 966 – 982.
- [8] Meidan Y, Lerner B, Rabinowitz G, et al. Cycle-time key factor identification and prediction in semiconductor manufacturing using machine learning and data mining [J]. IEEE Transactions on Semiconductor Manufacturing, 2012, 24(2):237 - 248.
- [9] 彭亚珍.基于 DBR 的多重人制造系统调度方法研究[D]. 北京:北京化工大学工学硕士论文,2012,7. Peng Yazhen. Scheduling method based on DBR for re-entrant manufacturing system[D]. Beijing: Dissertation for the Master Degree in Engineering of Beijing University of Chemical Technology, 2012, 7.
- [10] 乔非,马玉敏,李莉,等,基于分层瓶颈分析的多重入制 造系统调度方法[J]. 计算机集成制造系统, 2010, 16(4): 855 - 860.
 - Qiao Fei, Ma Yumin, Li Li et al. Multi-reentrant manufacturing system scheduling based on layered bottleneck analysis [J]. Computer Integrated Manufacturing System, 2010, 16 (4):855 – 860. (in Chinese)
- [11] Koh S G, Bulfin R L. Comparison of DBR with CONWIP in an unbalanced production line with three stations [J]. International Journal of Production Research, 2004, 42 (2): 391 -

404.

- [12] Wu H H, Yeh M L. A DBR scheduling method for manufacturing environments with bottleneck re-entrant flows[J]. International Journal of Production Research, 2006, 44(5): 883 902.
- [13] Chen C. A heuristic model for justifying the acceptance of rush orders in multi-plants manufacturing supply chain with outsourcing[C].8th IEEE International Conference on Industrial Informatics, 2010, 607 611.
- [14] Wang C, Liu G, Chen M, et al. A supply chain network sys-

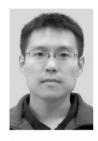
- tem optimum model for rush orders production decision in global manufacturing [C]. 8th IEEE International Conference on Industrial Informatics, 2010, 617 622.
- [15] Jang J R. ANFIS: adaptive-network-based fuzzy inference system[J]. IEEE Transactions on Systems, Man and Cybernetics, 1993,23(3):665 685.
- [16] Cao Z C, Zhao H D, Wang Y J. ANFIS and SA Based Approach to Prediction, Scheduling, and Performance Evaluation for Semiconductor Wafer Fabrication [J]. Chinese Journal of Electronics, 2013, 22(1):25 30.





曹政才 男,博士,教授/博导,主要研究方向为复杂生产过程优化调度,系统工程和智能控制.

E-mail: gifteze@163.com.



李 博 男,硕士,主要研究方向为复杂过程建模、优化与控制等.

E-mail: oberhere@126.com



彭亚珍 男,硕士,主要研究方向为复杂过程建模、优化与控制等.

E-mail: pengyazhen@163.com



刘 民 男,博士,教授/博导,主要研究方向为复杂生产过程优化调度,复杂生产过程或装备智能检测和智能操作优化,智能优化方法及其在复杂制造系统建模、调度与优化中的应用.E-mail;lium@tsinghua.edu.cn