基于分层分块结构的流程工业过程运行状态评价及非优原因追溯

邹筱瑜 [1] 王福利 [1] 常玉清 [1] 王 敏 [1] 蔡庆宏 [2]

摘 要 过程运行状态评价是指在安全生产的前提下, 对过程运行性能优劣程度做出评判, 并对非优运行状态追溯原因, 以指导操作人员进行生产调整. 针对含不确定性的流程工业过程运行状态评价问题, 本文提出分层分块评价结构. 同时, 考虑到实际生产过程中大量存在的不确定性信息, 采用粗糙集对每个子模型进行建模, 建立过程变量、工艺指标与综合评价指标间的关系. 最后, 将所提方法应用于一个金湿法冶金过程中, 验证所提方法有效性.

关键词 过程运行状态评价, 非优原因追溯, 流程工业, 粗糙集, 湿法冶金

引用格式 邹筱瑜, 王福利, 常玉清, 王敏, 蔡庆宏. 基于分层分块结构的流程工业过程运行状态评价及非优原因追溯. 自动化学报, 2019, **45**(2): 315−324

DOI 10.16383/j.aas.2017.c170159

Plant-wide Process Operating Performance Assessment and Non-optimal Cause Identification Based on Hierarchical Multi-block Structure

ZOU Xiao-Yu[1] WANG Fu-Li[1] CHANG Yu-Qing[1] WANG Min[1] CAI Qing-Hong[2]

Abstract Process operating performance assessment may determine the operating optimality degree and identify the cause of non-optimal operating performance, thus providing guidance for production adjustment. For the plant-wide process uncertainty, a novel hierarchical multi-block assessment framework is proposed in this paper. Due to the widely existing uncertainties in the plant-wide processes, the rough set is utilized to build an assessment model of each sub-block. The relations between process variables, technical indices and the comprehensive assessment index are extracted. Finally, the proposed method is applied to a gold hydrometallurgy process to illustrate its validity and feasibility.

Key words Process operating performance assessment, non-optimal cause identification, plant-wide process, rough set theory, gold hydrometallurgy process

Citation Zou Xiao-Yu, Wang Fu-Li, Chang Yu-Qing, Wang Min, Cai Qing-Hong. Plant-wide process operating performance assessment and non-optimal cause identification based on hierarchical multi-block structure. *Acta Automatica Sinica*, 2019, **45**(2): 315−324

过程运行状态评价指在过程安全、正常运行的前提下, 进一步对过程运行性能优劣做出评价, 并对非优原因进行追溯, 为操作人员提供合理可靠的操作指导建议. 过程运行状态优劣评价指标种类众多, 其中, 综合经济指标 (Comprehensive economic index, CEI) 是最常用的评价指标之一, CEI 将物耗、能耗、产品质量等指标融合为经济效益的综合指标. 有效的运行状态评价方法有助于保证过程 CEI 最大化. 因此, 近年来对运行状态评价方法的研究越来越多.

性能评价方法已广泛应用于各个领域[1−4]. 层次分析法 (Analytic hierarchy process, AHP) 是一种研究初期盛行的方法, AHP 以指标的拓扑结构为基础, 已广泛应用于环境和风险评价中[5−7]. 过程危害分析法 (Process hazards analysis, PHA) 和专家系统是基于过程知识的方法, 已应用于风险和状态评价中[8−9]. 灰色关联分析法 (Grey relational analysis, GRA) 是一种处理小样本情况的方法, 曾应用于环境和经济效益的评价[10]. 作为人工智能方法发展的一个重要成果, 人工神经网络 (Artificial neural network, ANN) 拥有强大的学习能力和非线性处理能力. 基于 ANN 的评价方法得到广泛应用, 例如临界状态评估等[11−12]. 其他常用评价方法还包

收稿日期 2017-03-27 录用日期 2017-08-17
Manuscript received March 27, 2017; accepted August 17, 2017
国家自然科学基金 (61533007, 61374146, 61174130) 资助
Supported by National Natural Science Foundation of China (61533007, 61374146, 61174130)
本文责任编委 苏宏业
Recommended by Associate Editor SU Hong-Ye
1. 东北大学信息科学与工程学院 沈阳 110819 2. 国家电网辽宁省电力有限公司 沈阳 110819
1. College of Information Science and Engineering, Northeastern University, Shenyang 110819 2. Liaoning Electric Power Co. LTD., State Grid, Shenyang 110819

括信息理论、动态概率理论等, 已应用于结构破坏评价[13]、空气质量评价[14] 等领域. 但上述典型评价方法并没有针对过程运行状态优劣程度在线评价进行深入研究.

过程运行状态在线评价是近几年诞生的一个新兴研究方向. 对于数据呈单峰分布的过程, Liu 等针对线性和非线性的情况, 分别提出了基于 Total projection to latent structure (T-PLS)[15] 和 Kernel total projection to latent structure (KT-PLS)[16] 的评价方法. 对于数据呈多峰分布的过程, Liu 等提出了基于高斯混合模型回归 (Gaussian mixture model regression, GMR)[17] 的评价方法. 此类方法虽能实现过程运行状态在线评价, 却过度依赖于数据质量, 要求数据测量准确、实时, 否则, 会使评价准确度和精度大大降低, 甚至造成评价模型失效.

在实际生产过程中, 存在大量无法准确、实时测量的变量, 变量取值可能来自于粗糙的测量、离线化验、专家定性估计等, 即不确定信息在实际流程工业过程中广泛存在. 此类不确定性主要来源于恶劣的生产环境和落后的测量技术. Pawlak 提出的粗糙集 (Rough set, RS) 理论是一种在不确定性存在的前提下进行推理的方法[18−19], 现已广泛应用于安全性评价和风险评价等领域[20−21], 但还未有研究将 RS 应用于含不确定性的过程运行状态最优性在线评价中.

当过程运行于非优运行状态时, 需进行原因追溯. 传统的非优原因追溯方法多基于确定性信息. 基于指标分解的非优原因追溯方法是最常用的方法, Liu 等[15] 和 Zou 等[22] 将优性评价指标分解为与变量相关的单项式和作为变量对优性指标的贡献和. 超出历史最优运行状态时贡献范围的变量, 被认定为非优原因变量. 只有优性指标可被分解为独立的变量贡献时, 此方法才有效. Liu 等[16−17] 将优性指标对每个变量的偏导数值作为相应变量的贡献, 并将超出最优运行状态时贡献范围的变量作为非优原因变量. 该方法仅适用于优性评价指标为连续可导函数的情况. 当过程存在不确定信息时, 上述两类方法都难以直接应用.

本文旨在解决含不确定信息的流程工业过程运行状态优性评价和非优原因追溯问题. 流程工业过程含有多生产单元、多变量等特点, 同一生产单元内变量耦合程度高, 单元间变量耦合程度低, 整体变量规模庞大. 若将传统评价方法直接应用于流程工业过程中, 可能存在以下三个主要问题: 1) 问题规模庞大, 模型复杂, 难以准确建立; 2) 过程变量并不直接影响最终评价指标, 而是通过一系列中间工艺指标, 逐步逐层影响综合评价指标, 因此, 难以直接提取过程变量与最终评价指标的相关性; 3) 在非优情况下, 非优原因变量难以快速定位, 模型解释性差. 考虑到实际复杂流程工业过程特点和数据特性, 本文提出基于 RS 的分层分块模型, 应用于过程运行状态最优性在线评价和非优原因追溯. 根据流程工业过程管理方向和生产流程进行纵向分层和横向分块, 用 RS 对每一个子块分别进行建模. 在分层分块模型的基础上, 进行评价和非优原因追溯. 本文所提方法考虑到不确定信息, 合理简化问题规模, 兼顾子块内和子块间的信息、不同层次指标之间的关系, 提高评价精度和解释性.

1　分层分块评价模型的建立

针对流程工业过程多变量、多生产单元、含不确定性、有不同层次指标的特点, 提出一种基于 RS 的分层分块模型 (RS based hierarchical multi-block model, RSHMM), 进行运行状态评价, 简化问题规模, 提高模型解释性.

1.1　分层分块结构的提出

为了简化运行状态评价问题的规模, 使得运行状态评价模型具有更强的物理意义, 根据工业过程流程、工艺特性和工艺指标等级的分布, 将一个流程工业过程划分为不同层次, 每一层划分为不同子块. 管理层次的细致程度决定了层次的数目, 过程规模和工艺特性决定了子块的数目. 本文采用的单元层、功能区层和全流程层的划分方法是一种常见的流程工业过程划分方法, 但不是唯一的划分方法. 一个流程工业过程的层次和子块的划分如图 1 所示. 纵向, 划分为全流程层、功能区层和单元层; 横向, 功能区层划分为 M 个功能区, 第 m 个功能区划分为 N_m 个运行单元. 全流程包含整个流程工业生产过程; 一个功能区包含独立完成一个生产功能的所有生产单元; 一个生产单元包含联系紧密的一系列设备, 同一生产单元内变量强耦合. 一个单元模型的输入为过程变量, 输出为单元层工艺指标; 在建立单元指标与单元内过程变量的模型时, 提取了与单元指标有关的变量相关性特征. 一个功能区模型的输入为体现单元特性的所有单元层工艺指标, 输出为功能区层工艺指标. 在建立功能区指标与单元指标的模型时, 提取了与相应功能区指标有关的单元指标之间的相关特征、耦合关系. 同理, 全流程模型的输入为各个功能区层工艺指标, 输出为全流程评价指标. 在建立全流程评价指标与功能区层工艺指标的关系时, 提取了与全流程评价指标有关的功能区指标之间的相关特征、耦合关系. 同时, 也相当于提取了与全流程指标相关的单元子块间的相关关系. 工艺指标一般包

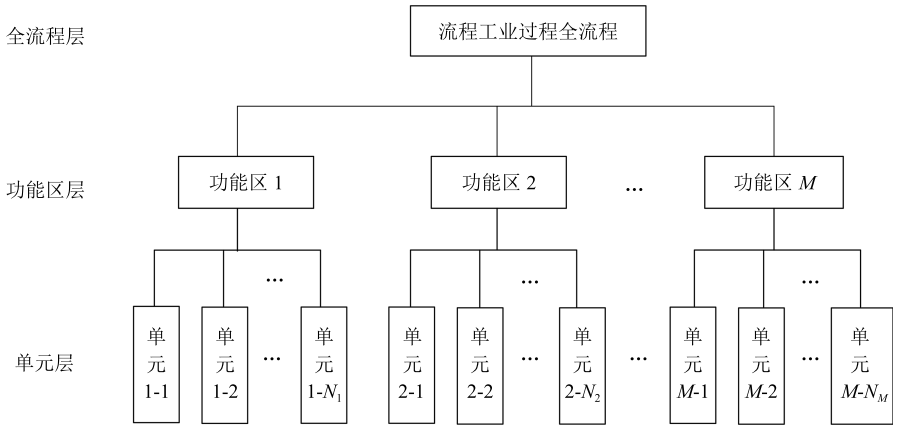


图 1 分层分块结构示意图

Fig. 1 Hierarchical multi-block structure

括质量指标和消耗指标. 在评价过程中, 不能单独通过一个子块的工艺指标判断该子块的优劣程度, 但可建立低一层工艺指标状态与高一层指标状态的关系, 最终得到全流程指标和相关工艺指标的关系. 同一模型内变量/指标之间相关性相对较强, 不同模型间变量/指标之间相关性相对较弱.

1.2 基于 RSHMM 的离线建模

在完成分层分块结构的建立之后, 由于不确定信息的存在, 采用 RS 对每一个子模型进行建模, 建立输入与输出之间的对应关系.

1.2.1 RS 理论简介

RS 是 Pawlak 提出的一种在信息系统内表达不确定信息的方法[18−19, 23]. 假设 U 为目标的非空有限集合, 称为论域, X 为论域的一个子集 (概念), A 是一个有限的属性集合, R 是 A 的一个子集. 对于论域中的一个元素 x, 令 $[x]_R$ 是一个包含 x 的集合, 其中, $[x]_R$ 中的元素在关系 R 上都相同. $[x]_R$ 称为 x 在关系 R 上的等价类. 对于论域的任何子集 X, X 的上、下近似分别定义为

$$\begin{aligned} R(X) &= \{x \in U | [x]_R \subseteq X\} \\ \bar{R}(X) &= \{x \in U | [x]_R \cap X \neq \emptyset\} \end{aligned} \tag{1}$$

其中, $R(X)$ 包含 U 中所有可以确定属于 X 的元素, $\bar{R}(X)$ 包含 U 中所有可能属于 X 的元素. X 的 R- 边界域定义为

$$BN_R(X) = \bar{R}(X) - R(X) \tag{2}$$

如果 X 的边界域为空, 即 $BN_R(X) = \emptyset$, 则称 X 是 R 可定义的, 否则称 X 是 R 不可定义的. R 可定义集称作 R- 精确集, R 不可定义集也称作 R-Rough 集, 在不发生混淆的情况下, 简称 Rough 集.

RS 的约简和推理在决策表的基础上进行. 决策表的每一列表示一个属性, 每个属性的取值被划分为若干离散状态. 通常, 属性可分为条件属性和决策属性. 决策表每一行代表论域中的一个元素和一种推理规则. 规则以 IF (条件), THEN (决策) 的形式表达, 来实现推理决策. RS 理论参见文献 [23].

1.2.2 离线建模

在实际生产中, 一部分变量可以实时、定量测量, 还有一部分变量无法准确、实时测量, 变量取值可能来自于粗糙的测量、离线化验、专家定性估计等. 本文旨在解决存在大量不确定信息的流程工业过程运行状态评价问题.

在建立如图 1 的分层分块结构基础上, 用 RS 建立每个子模型的输入和输出之间的关系. 采集充足的、覆盖所有运行状态等级的历史数据, 以分层分块结构为基础, 自顶向下, 确定各子模型的决策属性和条件属性, 组织决策表. 每个子模型的输入输出, 如图 2 所示. 离线建模数据包括全流程评价指标 CEI、功能区层评价指标、单元层评价指标、各单元内过程变量和表征生产工况的变量. 在全流程层, 以全流程运行状态综合评价指标 CEI 为决策属性, 以功能区层工艺指标 $\tilde{\boldsymbol{Z}} = [\boldsymbol{Z}^1, \boldsymbol{Z}^2, \cdots, \boldsymbol{Z}^M]$ 和代表过程生产工况的变量 $\boldsymbol{C}$ 为条件属性, 组织决策表, 其中, $\boldsymbol{Z}^m = [\boldsymbol{Z}_1^m, \boldsymbol{Z}_2^m, \cdots, \boldsymbol{Z}_{I_m}^m]$ $(m = 1, 2, \cdots, M)$ 为功能区 m 的 I_m 个指标, $\boldsymbol{Z}_i^m \in \mathbf{R}^{D\times 1}$ 为第 i 个指标的取值, D 为历史样本数目. 过程运行状态可根据 CEI 的水平划分为不同等级, CEI 越高, 反映出运行状态等级越好. 在功能区层, 针对第 m 个功能区, 以该功能区工艺指标分别为决策属性, 以单元层工艺指标 $\tilde{\boldsymbol{Y}}^m = [\boldsymbol{Y}^{m,1}, \boldsymbol{Y}^{m,2}, \cdots, \boldsymbol{Y}^{m,N_m}]$ 和 $\boldsymbol{C}$ 为条件属性, 组织第 m 个功能区的决策表, 其中, $\boldsymbol{Y}^{m,n} = [\boldsymbol{Y}_1^{m,n}, \boldsymbol{Y}_2^{m,n}, \cdots, \boldsymbol{Y}_{I_{m,n}}^{m,n}]$ $(n = 1, 2, \cdots, N_m)$ 为第

m 个功能区第 n 个单元的 $I_{m,n}$ 个指标, $\boldsymbol{Y}_i^m \in \mathbf{R}^{D\times 1}$ 为第 i 个指标的取值. 在单元层, 针对第 m 个功能区第 n 个单元, 分别以该单元工艺指标为决策属性, 以该单元内过程变量 $\tilde{\boldsymbol{X}}^{m,n}$ 和 $\boldsymbol{C}$ 为条件属性, 组织该单元的决策表, 其中, $\boldsymbol{Z}_i^m \in \mathbf{R}^{D\times J_{m,n}}$, $J_{m,n}$ 为第 m 个功能区第 n 个单元包含的过程变量个数. 不同生产工况下, 过程对同样操作的响应特性不同, 运行状态性能的评价指标所能达到的水平不同, 性能评价标准也不同. 因此, 需要在每一层次每一子块的输入中加入反应过程生产工况的信息 $\boldsymbol{C}$. 如果过程运行于单一工况, 那么不需要进行工况区分, $\boldsymbol{C}$ 为空. 建立各子模型的决策表之后, 将过程变量、工艺指标和评价指标进行离散化, 即把无限空间中无限的个体映射到有限的空间中去, 以此提高算法的时空效率. 变量和指标的状态划分可采用传统的等距离离散化方法[23] 或聚类方法[24], 也可借助专家经验[25] 进行划分. 基于分层分块结构的评价模型, 自顶向下进行离散化. 离散化是应对过程信息不确定性的有效手段. 针对定量测量的变量, 将变量取值离散化为不同状态等级, 并用一系列正整数对不同等级进行顺次编号. 针对定性估计的变量, 将每一种定性状态作为一个变量状态等级, 并进行编号. 值得注意的是, 本文所指变量状态等级只与变量幅值大小有关, 与变量优劣无关, 只有全流程运行状态等级才与性能优劣有关; 后文中, 假设所有变量均已合理离散化.

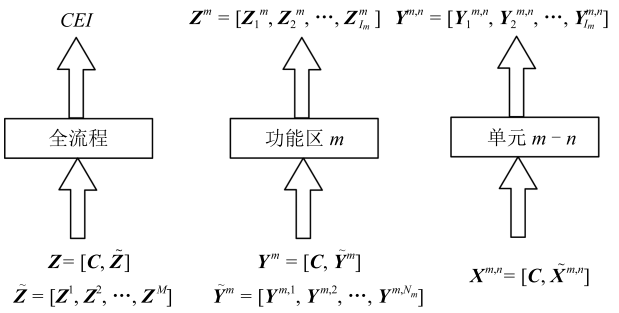


图 2　离线模型输入输出示意图

Fig. 2　The diagram of input and output for the offline model

在组织决策表并离散化之后, 需要对每一张决策表进行约简, 即在保持分类能力不变的前提下, 删除对决策没有影响的条件属性[26−27]. 常用的属性约简方法有一般约简算法、基于差别矩阵和逻辑运算的属性约简算法、归纳属性约简算法等.

离线建模最后的操作是提取决策规则. 规则库中决策规则为 IF (条件), THEN (决策) 的形式, 即启发式规则. 条件为各条件属性状态取逻辑和的形式, 决策为相应决策属性的状态. 决策规则是进行分层分块评价和追溯的基础.

2　基于 RSHMM 的流程工业过程运行状态评价

2.1　在线评价

通常, 流程工业过程中间工艺指标和综合评价指标难以直接在线获得. 因此, 本文利用建立的分层分块模型, 自底向上逐步推理出工艺指标和综合评价指标的状态. 1) 通过过程变量的测量值, 利用单元层各 RS 模型推理出单元层各工艺指标的状态; 2) 以单元层工艺指标状态为基础, 用功能区层 RS 模型推理出功能区层各工艺指标的状态; 3) 以功能区层工艺指标状态为基础, 用全流程层 RS 模型推理全流程综合评价指标 CEI 的状态. 最终运行状态等级评价策略以 CEI 在一定时间窗口内的状态为基础, 进行进一步判断.

获得时刻 t 的在线数据后, 用与离线建模相同的离散化方法, 得到离散化后的数据 $\boldsymbol{x}_t = [\boldsymbol{c}_t, \tilde{\boldsymbol{x}}_t] \in \mathbf{R}^{1\times J}$, 其中, $\boldsymbol{c}_t$ 表示过程生产工况, $\tilde{\boldsymbol{x}}_t$ 表示各子块内包含的变量, J 为变量维数.

令 $\boldsymbol{x}_t^{m,n} = [\boldsymbol{c}_t, \tilde{\boldsymbol{x}}_t^{m,n}]$ 为推理第 m 个功能区中第 n 个单元工艺指标的基础, 其中, $\tilde{\boldsymbol{x}}_t^{m,n} \in \mathbf{R}^{1\times J^{m,n}}$ 包含该单元内的过程变量, $J^{m,n}$ 为过程变量个数. 根据单元层工艺指标的 RS 模型, 以 $\boldsymbol{x}_t^{m,n}$ 为条件属性, 推理出单元层各工艺指标 $\boldsymbol{y}_t^{m,n} = [y_{t,1}^{m,n}, y_{t,2}^{m,n}, \cdots, y_{t,I_{m,n}}^{m,n}]$ 的状态.

令 $\boldsymbol{y}_t^m = [\boldsymbol{c}_t, \tilde{\boldsymbol{y}}_t^m]$ 为推理第 m 个功能区工艺指标的基础, 其中, $\tilde{\boldsymbol{y}}_t^m = [\boldsymbol{y}_t^{m,1}, \boldsymbol{y}_t^{m,2}, \cdots, \boldsymbol{y}_t^{m,N_m}]$. 根据功能区层工艺指标的 RS 模型, 以 $\boldsymbol{y}_t^m$ 为条件属性, 推理出功能区层各工艺指标 $\boldsymbol{z}_t^m = [z_{t,1}^m, z_{t,2}^m, \cdots, z_{t,I_m}^m]$ 的状态.

令 $\boldsymbol{z}_t = [\boldsymbol{c}_t, \tilde{\boldsymbol{z}}_t]$ 为推理全流程综合评价指标 CEI 的基础, 其中, $\tilde{\boldsymbol{z}}_t = [\boldsymbol{z}_t^1, \boldsymbol{z}_t^2, \cdots, \boldsymbol{z}_t^M]$. 根据全流程 RS 评价模型, 由 $\boldsymbol{z}_t$ 推理出此时的评价指标, 记为 CEI_t.

为降低评价对噪声的敏感度, 时刻 t 的等级判断不能以单个时刻的 CEI_t 为标准. 引入长度为 H 的滑动窗口, 以 t 时刻滑动窗口 $\{CEI_{t-H+1}, CEI_{t-H+2}, \cdots, CEI_t\}$ 内, 出现频率最高的 CEI 状态等级作为此时运行状态等级的评价结果, 记为 Gd_t. 若出现频率最高的 CEI 状态等级不止一个, 则保持前一时刻评价结果不变. 运用此评价规则, 针对一个存在 L 个运行状态等级的过程, 其初始状态的确定, 最多需要 λ 个时刻的评价, 其中, $\lambda = \max\{L+1, H\}$.

在确定初始运行状态等级后, 具体在线评价步骤总结如下:

步骤 1. 获取 t 时刻在线数据并离散化得 $\boldsymbol{x}_t =$

$[\boldsymbol{c}_t, \tilde{\boldsymbol{x}}_t]$;

步骤 2. 构造推理第 m 个功能区中第 n 个单元工艺指标的基础 $\boldsymbol{x}_t^{m,n} = [\boldsymbol{c}_t, \tilde{\boldsymbol{x}}_t^{m,n}]$;

步骤 3. 根据单元层工艺指标的 RS 模型, 由 $\boldsymbol{x}_t^{m,n}$ 推理 $\boldsymbol{y}_t^{m,n} = [y_{t,1}^{m,n}, y_{t,2}^{m,n}, \cdots, y_{t,I_{m,n}}^{m,n}]$;

步骤 4. 构造推理第 m 个功能区工艺指标的基础 $\boldsymbol{y}_t^m = [\boldsymbol{c}_t, \tilde{\boldsymbol{y}}_t^m]$, 其中, $\tilde{\boldsymbol{y}}_t^m = [\boldsymbol{y}_t^{m,1}, \boldsymbol{y}_t^{m,2}, \cdots, \boldsymbol{y}_t^{m,N_m}]$;

步骤 5. 根据功能区层工艺指标的 RS 模型, 由 $\boldsymbol{y}_t^m$ 推理 $\boldsymbol{z}_t^m = [z_{t,1}^m, z_{t,2}^m, \cdots, z_{t,I_m}^m]$;

步骤 6. 构造推理全流程综合评价指标 CEI 的基础 $\boldsymbol{z}_t = [\boldsymbol{c}_t, \tilde{\boldsymbol{z}}_t]$, 其中, $\tilde{\boldsymbol{z}}_t = [\boldsymbol{z}_t^1, \boldsymbol{z}_t^2, \cdots, \boldsymbol{z}_t^M]$;

步骤 7. 根据全流程 RS 评价模型, 由 $\boldsymbol{z}_t$ 推理出 CEI_t;

步骤 8. 以 t 时刻滑动窗口 $\{CEI_{t-H+1}, CEI_{t-H+2}, \cdots, CEI_t\}$ 内, 出现频率最高的 CEI 状态等级作为此时运行状态等级的评价结果, 记为 Gd_t. 若出现频率最高的 CEI 状态等级不止一个, 则保持前一时刻评价结果不变.

2.2 非优原因追溯

非优原因追溯的目的是在过程运行状态处于非优等级时, 找到导致非优状态的原因, 为操作人员提供生产指导. 为查找真正的非优原因变量, 提供使生产调整代价最小的最优生产条件作为参考, 本文提出一种基于匹配度的非优原因追溯方法, 查找非优原因, 提供调整方向, 保证调整的有效性和可行性. 所提方法考虑到了最优生产状态与当前生产情况的差异度、变量调整的幅度、变量调整的难易程度等.

在分层分块的基础上, 自顶向下, 进行查找. 本文通过衡量当前非优数据与最优运行状态规则库中数据的匹配程度, 选取与当前非优数据匹配度最大的优规则作为参考, 与此优参考数据差异度较大的属性为非优的属性. 对于一个非优的指标, 需要继续以相同方法向下查找更深层次的非优原因. 直到查找到单元层非优的过程变量为止, 得到非优原因.

一次查找中, 规则匹配的过程如图 3 所示, 其中, $a_1a_2a_3\cdots a_{J'}$ 和 $a_1^*a_2^*a_3^*\cdots a_{J'}^*$ 分别表示当前的非优数据在各属性的状态取值和与之匹配度最大的优规则在各属性的状态取值, J' 为条件属性个数. 用匹配度公式衡量当前数据与优规则的匹配情况. 匹配度选取的准则为: 1) 两个数据中相同的属性越多, 匹配度越大; 2) 两个数据中属性等级的级差越大, 匹配度越小; 3) 两个数据中不同的属性调整越容易, 匹配度越大.

定义 $\boldsymbol{a} = [a_1, a_2, \cdots, a_{J'}]$ 和第 k 条优规则 $\boldsymbol{a}_k = [a_{k,1}, a_{k,2}, \cdots, a_{k,J'}]$ 的匹配度为

$$\gamma(\boldsymbol{a}, \boldsymbol{a}_k) = 1 - \sum_{j=1}^{J'} \left(\frac{|a_j - a_{k,j}|}{A_j - 1} \times \beta_j \right) \tag{3}$$

其中, $|a_j - a_{k,j}|$ 为 a_j 和 $a_{k,j}$ 的等级差, A_j 为属性 j 的等级总数, β_j 为属性 j 的权重, 权重越小, 相应属性可操作性越强, 且满足 $\sum_{j=1}^{J'} \beta_j = 1$, 利用式 (3), 可在最优规则库中选取与 $\boldsymbol{a}$ 匹配度最高的规则, 记为 $\boldsymbol{a}^*$. 然后利用差异度, 确定 $\boldsymbol{a}$ 确定 $\boldsymbol{a}$ 中非优的属性. 定义 $\boldsymbol{a}$ 和 $\boldsymbol{a}^*$ 中第 j 个属性的差异度为

$$d(a_j, a_j^*) = \frac{\left|a_j - a_j^*\right|}{(A_j - 1) \sum\limits_{j'=1}^{J'} \frac{\left|a_{j'} - a_{j'}^*\right|}{A_{j'} - 1}} \tag{4}$$

显然, $d(a_j, a_j^*) \in [0, 1]$, $\sum_{j'=1}^{J'} d(a_{j'}, a_{j'}^*) = 1$. $d(a_j, a_j^*)$ 较大的属性为非优属性.

基于 RSHMM 的非优原因追溯如图 4 所示. 在全流程层, 需要查找 $\boldsymbol{z}_t$ 中导致非优的功能区层指标. 以全流程层 RS 模型为基础, 将决策属性 CEI 为优的数据构成全流程层的最优规则库. 计算 $\boldsymbol{z}_t$ 与最优规则库中每条规则的匹配度, 得到与之匹配度最大

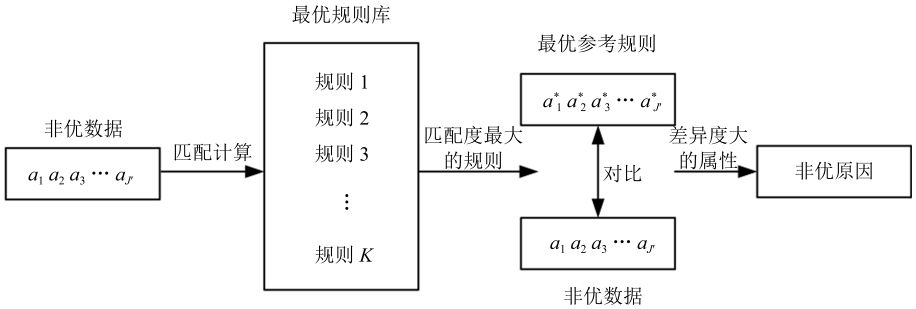


图 3 非优原因追溯中的规则匹配示意图

Fig. 3 Schematic diagram of the rule matching in non-optimal cause identification

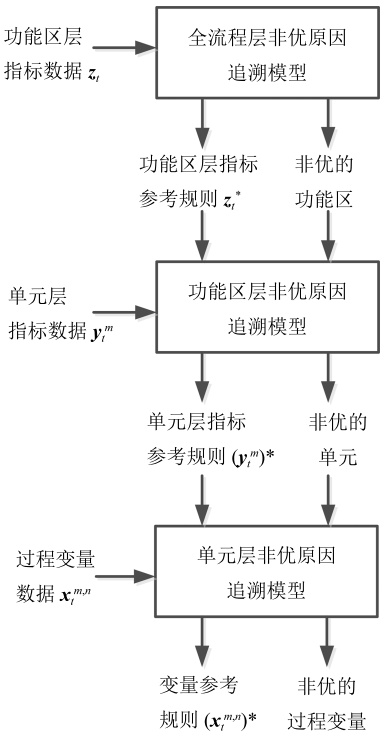


图 4　基于 RSHMM 的非优原因追溯示意图

Fig. 4　Schematic diagram of non-optimal cause identification based on RSHMM

的优参考规则, 记为 $\boldsymbol{z}_t^*$. 对比 $\boldsymbol{z}_t$ 和 $\boldsymbol{z}_t^*$ 中的每个属性状态取值的差异, 差异度大的为非优的属性, 需要进一步查找其非优原因. 在获得 $\boldsymbol{z}_t^*$ 的同时, 对应的单元层指标和过程变量也可以相应获得, 这些数据构成新的优规则库, 是下一层追溯的基础.

在非优的功能区内, 对于一个非优的指标 $z_{t,i}^m$, 需要查找 $\boldsymbol{y}_t^m$ 中导致非优的单元层指标. 以第 m 个功能区第 i 个工艺指标的 RS 模型为对象, 计算 $\boldsymbol{y}_t^m$ 与最优规则库中每条规则的匹配度, 得到与之匹配度最大的优参考规则, 记为 $(\boldsymbol{y}_t^m)^*$. 对比 $\boldsymbol{y}_t^m$ 和 $(\boldsymbol{y}_t^m)^*$ 中的每个属性状态取值的差异, 差异度大的为非优的属性, 需要进一步查找其非优原因. 此时, 优规则库进一步缩小为功能区层指标为 $\boldsymbol{z}_t^*$、单元层指标为 $(\boldsymbol{y}_t^m)^*$ 的规则, 作为下一层追溯的基础.

在非优的单元内, 对于一个非优的指标 $y_{t,i}^{m,n}$, 需要查找 $\boldsymbol{x}_t^{m,n}$ 中导致非优的过程变量. 类似地, 以第 m 个功能区第 n 个单元第 i 个工艺指标的 RS 模型为对象, 通过计算匹配度, 得到最优规则库中与 $\boldsymbol{x}_t^{m,n}$ 匹配度最大的优参考规则 $(\boldsymbol{x}_t^{m,n})^*$. 对比 $\boldsymbol{x}_t^{m,n}$ 和 $(\boldsymbol{x}_t^{m,n})^*$ 中的每个变量状态取值的差异, 差异度大的为非优原因变量. 如果出现针对同一单元的多种追溯结果, 则将所有追溯结果进行输出, 由操作人员或领域专家结合过程知识进行进一步判断.

非优原因变量的调整, 可参考 $(\boldsymbol{x}_t^{m,n})^*$ 中的状态. 最基本的调整策略有两种: 1) 将 $\boldsymbol{x}_t^{m,n}$ 和 $(\boldsymbol{x}_t^{m,n})^*$ 存在差异的变量, 按照差异度从大到小逐个进行调整; 2) 同时调整 $\boldsymbol{x}_t^{m,n}$ 和 $(\boldsymbol{x}_t^{m,n})^*$ 存在差异的变量. 第一种调整方法通常不需要等到所有变量都进行调整, 就可以恢复最优运行状态, 所需调节变量少, 但调整时间长; 第二种调整方法可以使生产过程迅速恢复到最优运行状态, 但是需要多个生产环节的相互配合. 具体调整策略的选择, 需根据过程特点和生产需求进行确定.

3　仿真分析

金湿法冶金过程是一个典型的流程工业过程. 与传统火法冶金不同, 金湿法冶金将矿石中固态的金转换为矿浆中液态的金, 再用锌粉进行置换, 在冶金效率和环境保护方面有较大优势. 但是, 金湿法冶金过程含有大量不确定性, 很多变量只能定性估计, 而不能定量测量. 因此, 本章以国内某金湿法冶金过程为背景, 验证所提方法的有效性和可行性. 该金湿法冶金工艺流程如图 5 所示, 过程包含三个功能区:

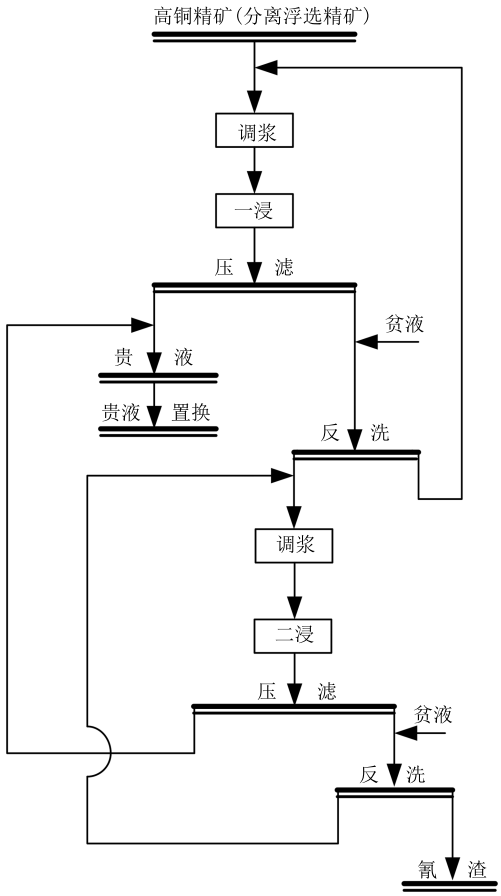


图 5　金湿法冶金工艺流程示意图

Fig. 5　The flow chart of the gold hydrometallurgy process production

浸出功能区、洗涤功能区和置换功能区. 浸出功能区包含一次浸出单元和二次浸出单元. 浸出单元用适当的溶剂处理矿石或精矿, 使黄金以金氰络合物离子形态进入溶液, 而脉石及其他杂质不溶解. 因此, 黄金和原矿物杂质以固–液相的形式区分开来. 洗涤功能区包含一次洗涤单元和二次洗涤单元. 压滤洗涤单元起到固液分离的作用, 将富含金氰络合物离子的贵液输送到置换功能区, 将固体杂质进行回收利用. 两浸两洗的工艺设置保证了黄金浸出率. 置换功能区利用锌粉与金氰络合物离子发生的置换反应, 将液相的黄金重新变为固相. 此功能区规模不大, 变量相关性强, 不需再细分为不同生产单元. 因此, 金湿法冶金过程的分层分块结构为: 全流程层是金湿法冶金全流程; 功能区层包含浸出、洗涤和置换功能区; 单元层包括浸出功能区中的第一次浸出和第二次浸出单元, 以及洗涤功能区中的第一次洗涤和第二次洗涤单元.

本节将所提方法应用于我们课题组开发的金湿法冶金半实物仿真平台中. 此仿真平台模拟了所研究的金湿法冶金生产过程. 经过长时间的实践、修正和完善, 可以较为准确地模拟该湿法冶金生产过程, 为实际生产决策提供参考. 经过深入研究, 选取与 CEI 密切相关的 15 个工艺指标和 34 个过程变量用于实现该过程的分层分块结构运行状态评价, 分别列于表 1 和表 2 中. 过程运行状态划分为优、中、差三个等级, 分别对应 CEI 高、中、低三个状态. 从半实物仿真平台采集三个等级的数据各 2 000 组作为建模数据, 每组数据包含生产条件、过程变量、单元层指标、功能区层指标和全流程评价指标 CEI 的

表 1 各层评价指标列表

Table 1 The assessment indices for each level

序号	指标名称	位置
1	一次浸出率	一次浸出单元层
2	一浸能耗	一次浸出单元层
3	一浸物耗	一次浸出单元层
4	二次浸出率	二次浸出单元层
5	二浸能耗	二次浸出单元层
6	二浸物耗	二次浸出单元层
7	一次洗涤率	一次洗涤单元层
8	二次洗涤率	二次洗涤单元层
9	总浸出率	浸出功能区层
10	总浸出能耗	浸出功能区层
11	总浸出物耗	浸出功能区层
12	总洗涤率	总洗涤率
13	置换率	置换功能区层
14	置换能耗	置换功能区层
15	置换物耗	置换功能区层
16	CEI	全流程层

表 2 湿法冶金过程变量表

Table 2 The variables of the gold hydrometallurgy process

序号	指标名称	位置
1	矿石来料量	生产条件
2	初始金品位	生产条件
3	矿石平均粒径	生产条件
4	调浆后矿浆浓度	一次浸出单元
5	调浆后矿浆流量	一次浸出单元
6	一浸前调浆水量	一次浸出单元
7	一浸 NaCN 添加量	一次浸出单元
8	一浸后金氰离子浓度	一次浸出单元
9	一浸空气流量	一次浸出单元
10	一浸溶氧量	一次浸出单元
11	一浸矿浆密度	一次浸出单元
12	一次浸出单元	二次浸出单元
13	二浸前矿浆流量	二次浸出单元
14	二浸前调浆水量	二次浸出单元
15	二浸 NaCN 添加量	二次浸出单元
16	二浸后金氰离子浓度	二次浸出单元
17	二浸空气流量	二次浸出单元
18	二浸溶氧量	二次浸出单元
19	二浸矿浆密度	二次浸出单元
20	一洗前矿浆浓度	一次压滤洗涤单元
21	一洗后金氰离子浓度	一次压滤洗涤单元
22	一洗后滤液流量	一次压滤洗涤单元
23	一洗后滤液流量	一次压滤洗涤单元
24	二洗前矿浆浓度	二次压滤洗涤单元
25	二洗后金氰离子浓度	二次压滤洗涤单元
26	二洗后滤液流量	二次压滤洗涤单元
27	二洗后滤饼流量	二次压滤洗涤单元
28	脱氧塔 1 压力	置换功能区
29	流入置换压滤机贵液流量	置换功能区
30	流入置换压滤机贵液金氰离子浓度	置换功能区
31	锌粉添加量	置换功能区
32	锌粉平均粒径	置换功能区
33	贫液流量	置换功能区
34	贫液金氰离子浓度	置换功能区

变量状态. 采集 300 组数据作为测试数据, 每组数据包括生产条件、单元层各子块所含变量的状态. 这 300 组测试数据模拟了运行状态等级从优变为中的生产过程: 前 150 组测试数据运行于优运行状态; 从第 151 组测试数据开始, 模拟矿浆来料量增多时, 未能及时增加二浸 NaCN 添加量, 导致运行状态从优变为中的过程. 选取滑动窗口长度 $H = 5$, 用本文所提 RSHMM 方法分别进行运行状态评价和非优原因追溯.

图 6 是基于本文所提 RSHMM 方法的全流程评价结果, 能较为准确地判断出运行状态等级的变化, 几个时刻的判断延迟主要来源于离散化步骤的信息损失. RSHMM 方法不仅能够得到全流程运行状态等级的评价结果, 还能得到各级过程指标的状态, 利于提高解释性, 快速定位导致非优运行状态的原因变量, 对生产调整提供清晰的指导. 非优原因追溯结果如图 7 所示, 所提方法能够准确定位非优原因变量, 并且深入展示非优运行状态产生的原因.

同时, 将该过程用传统 RS 方法进行评价. 经过反复实验证明, 基于 RSHMM 和 RS 评价方法的正确率与建模数据量有较大关系. 表 3 是在不同的建

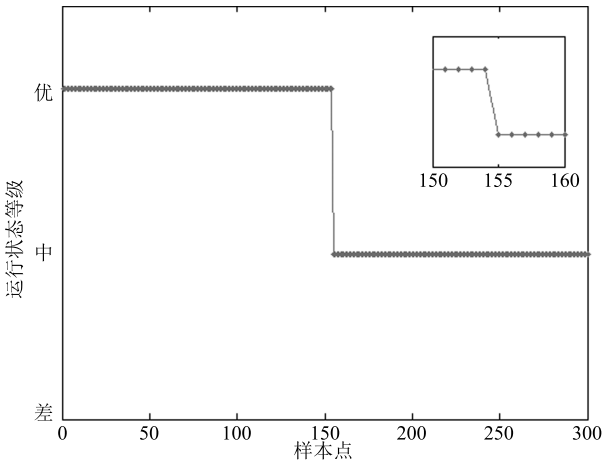


图 6　基于 RSHMM 的在线评价结果

Fig. 6　RSHMM based online assessment result

表 3　RSHMM 和 RS 评价正确率对比 (%)

Table 3　The assessment accuracy rate comparison of RSHMM and RS (%)

建模数据总量	RSHMM 的评价正确率	RS 的评价正确率
1 500	68.0	32.5
3 000	81.7	58.7
6 000	92.3	76.1
12 000	94.1	87.3
15 000	95.3	95.2

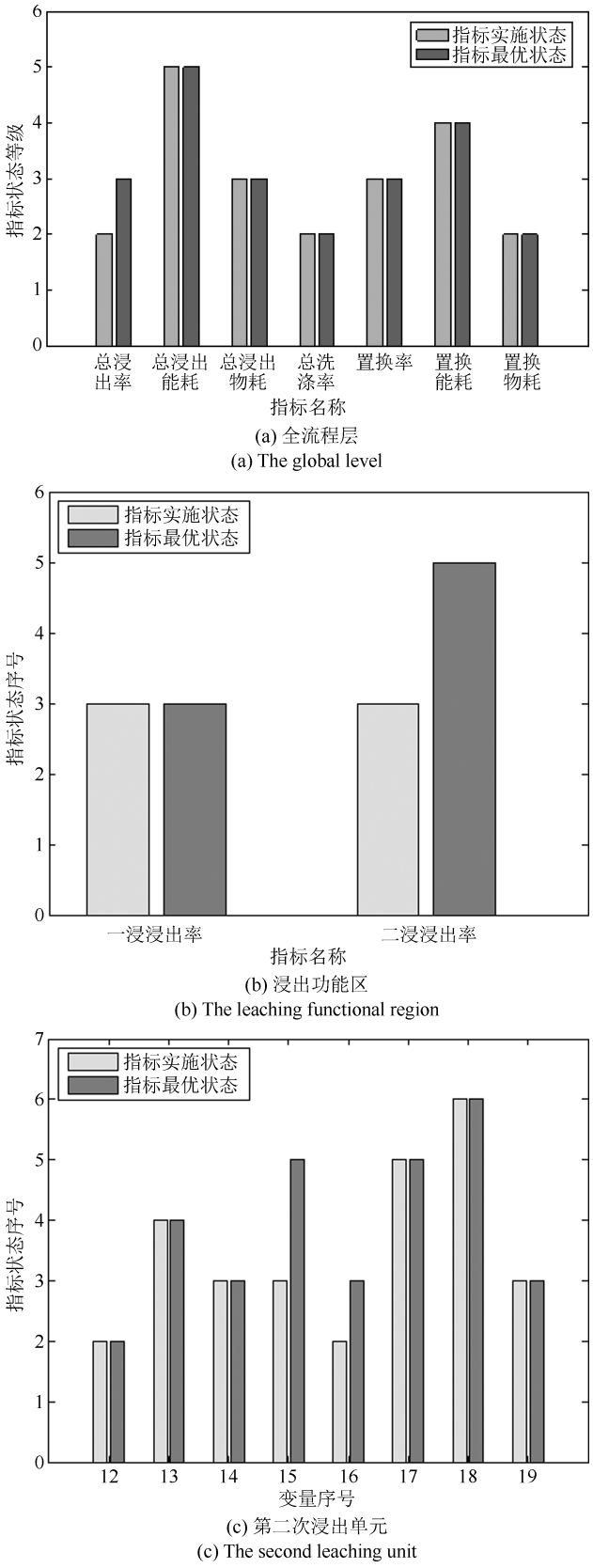


图 7　基于 RSHMM 的非优原因追溯结果

Fig. 7　RSHMM based non-optimal cause identification result

模数据量下, RSHMM 和 RS 评价的正确率. 可见, 当建模数据量少时, RSHMM 的正确率远大于 RS. 而当建模数据量大到一定程度时, 两种方法正确率差距较小. 用基于 RS 的方法进行追溯, 追溯结果虽然包含了正确的非优原因, 但同时也包含了其他因素, 追溯结果解释性差.

综上所述, 相比于传统的不分层评价方法, 本文所提方法有如下优势: 1) 对建模数据量的要求更低; 2) 解释性强, 展示各层次指标和变量的状态, 利于深入分析过程运行状态产生的原因; 3) 候选非优原因变量范围小, 更符合实际生产需求; 4) 可以清晰、准确、快速定位非优原因变量, 指导生产调整.

4 结论

针对含不确定信息的流程工业过程运行状态评价问题, 本文考虑到流程工业特点和数据特性, 提出一种基于分层分块结构的流程工业过程运行状态评价和非优原因追溯方法. 所提方法清晰直观地表达了流程工业工艺特性, 以及变量、指标之间的层次关系. 因此, 可以通过过程变量, 逐步推算综合评价指标的状态, 实现全流程的评价. 对于非优运行状态, 在分层分块结构的基础上, 利用匹配度公式, 逐步查找非优的指标和过程变量, 最终确定导致非优的原因变量, 提供生成指导. 最后, 将所提方法应用于某金湿法冶金过程, 取得了良好效果, 证明了所提方法的有效性、准确性和可解释性. 考虑到不确定信息, 各子块分别用 RS 建模, 但 RS 离散化过程中存在信息损失, 评价精度有限. 后续研究将会集中于提高运行状态评价精度.

References

1 Ang B W, Su B, Wang H. A spatial-temporal decomposition approach to performance assessment in energy and emissions. *Energy Economics*, 2016, **60**: 112−121

2 Sarri S, Zanni L, Popovic M, Le Boudec J Y, Paolone M. Performance assessment of linear state estimators using synchrophasor measurements. *IEEE Transactions on Instrumentation and Measurement*, 2016, **65**(3): 535−548

3 Zhao C H, Gao F R. Critical-to-fault-degradation variable analysis and direction extraction for online fault prognostic. *IEEE Transactions on Control Systems Technology*, 2017, **25**(3): 842−854

4 Gullo P, Elmegaard B, Cortella G. Energy and environmental performance assessment of R744 booster supermarket refrigeration systems operating in warm climates. *International Journal of Refrigeration*, 2016, **64**: 61−79

5 Ramanathan R. A note on the use of the analytic hierarchy process for environmental impact assessment. *Journal of Environmental Management*, 2001, **63**(1): 27−35

6 Handfield R, Walton S V, Sroufe R, Melnyk S A. Applying environmental criteria to supplier assessment: a study in the application of the analytical hierarchy process. *European Journal of Operational Research*, 2002, **141**(1): 70−87

7 Mustafa M A, Al-Bahar J F. Project risk assessment using the analytic hierarchy process. *IEEE Transactions on Engineering Management*, 1991, **38**(1): 46−52

8 Khan F I, Abbasi S A. Techniques and methodologies for risk analysis in chemical process industries. *Journal of Loss Prevention in the Process Industries*, 1998, **11**(4): 261−277

9 Saha T K, Purkait P. Investigation of an expert system for the condition assessment of transformer insulation based on dielectric response measurements. *IEEE Transactions on Power Delivery*, 2004, **19**(3): 1127−1134

10 Xu G, Yang Y P, Lu S Y, Li L, Song X N. Comprehensive evaluation of coal-fired power plants based on grey relational analysis and analytic hierarchy process. *Energy Policy*, 2011, **39**(5): 2343−2351

11 Guo L J, Gao J J, Yang J F, Kang J X. Criticality evaluation of petrochemical equipment based on fuzzy comprehensive evaluation and a BP neural network. *Journal of Loss Prevention in the Process Industries*, 2009, **22**(4): 469−476

12 Kumar S, Kaur T. Development of ANN based model for solar potential assessment using various meteorological parameters. *Energy Procedia*, 2016, **90**: 587−592

13 Meruane V, Ortiz-Bernardin A. Structural damage assessment using linear approximation with maximum entropy and transmissibility data. *Mechanical Systems and Signal Processing*, 2015, **54−55**: 210−223

14 Thunis P, Clappier A. Indicators to support the dynamic evaluation of air quality models. *Atmospheric Environment*, 2014, *98*: 402−409

15 Liu Y, Chang Y Q, Wang F L. Online process operating performance assessment and nonoptimal cause identification for industrial processes. *Journal of Process Control*, 2014, **24**(10): 1548−1555

16 Liu Y, Chang Y Q, Wang F L, Ma R C, Zhang H L. Complex process operating optimality assessment and nonoptimal cause identification using modified total kernel PLS. In: Proceedings of the 26th Chinese Control and Decision Conference. Changsha, China: IEEE, 2014. 1221−1227

17 Liu Y, Wang F L, Chang Y Q, Ma R C. Operating optimality assessment and nonoptimal cause identification for non-Gaussian multimode processes with transitions. *Chemical Engineering Science*, 2015, **137**: 106−118

18 Pawlak Z. Rough sets. *International Journal of Computer and Information Sciences*, 1982, **11**(5) 341−356

19 Pawlak Z. *Rough Sets: Theoretical Aspects of Reasoning about Data*. Dordrecht, Netherlands: Springer, 1991.

20 Zhong Z W, Mu L L. The integrated methodology of rough set theory and artificial neural network for safety assessment

on construction sites. In: Proceedings of the 2011 International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC). Hangzhou, China: IEEE, 2011. 105−109

21 Zhou J G, Bai T. Credit risk assessment using rough set theory and GA-based SVM. In: Proceedings of the 3rd International Conference on Grid and Pervasive Computing Workshops. Kunming, China: IEEE, 2008. 320−325

22 Zou X Y, Wang F L, Chang Y Q, Zhang B. Process operating performance optimality assessment and non-optimal cause identification under uncertainties. *Chemical Engineering Research and Design*, 2017, **120**: 348−359

23 Wang Guo-Yin. *Rough Set Theory and Knowledge Acquisition*. Xi′an: Xi′an Jiaotong University Press, 2001. (王国胤. Rough 集理论与知识获取. 西安: 西安交通大学出版社, 2001.)

24 Hartigan J A, Wong M A. Algorithm AS 136: a k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 1979, **28**(1): 100−108

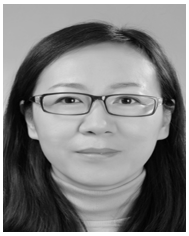
25 Wagstaff K, Cardie C, Rogers S, Schrödl S. Constrained k-means clustering with background knowledge. In: Proceedings of the 8th International Conference on Machine Learning. San Francisco, USA: Morgan Kaufmann Publishers Inc, 2001. 577−584

26 Xiao Di, Hu Shou-Song. Real rough set theory and attribute reduction. *Acta Automatica Sinica*, 2007, **33**(3): 253−258 (肖迪, 胡寿松. 实域粗糙集理论及属性约简. 自动化学报, 2007, **33**(3): 253−258)

27 Zhang Ying-Chun, Guo He. Level set image segmentation based on rough set and new energy formula. *Acta Automatica Sinica*, 2015, **41**(11): 1913−1925 (张迎春, 郭禾. 基于粗糙集和新能量公式的水平集图像分割. 自动化学报, 2015, **41**(11): 1913−1925)

邹筱瑜 东北大学信息科学与工程学院博士研究生. 主要研究方向为复杂工业过程运行状态评价与非优原因追溯.
E-mail: xiaoyuzou_neu@hotmail.com
(**ZOU Xiao-Yu** Ph. D. candidate at the College of Information Science and Engineering, Northeastern University. Her research interest covers complex process operating performance optimality assessment and non-optimal cause identification.)

王福利 东北大学信息科学与工程学院教授. 主要研究方向为复杂工业过程运行状态评价, 优化, 故障诊断. 本文通信作者. E-mail: wangfuli@ise.neu.edu.cn
(**WANG Fu-Li** Professor at the College of Information Science and Engineering, Northeastern University. His research interest covers operating performance optimality assessment, optimization, and fault diagnosis of complex system. Corresponding author of this paper.)



常玉清 东北大学信息科学与工程学院教授. 主要研究方向为复杂工业过程运行状态评价和故障诊断.
E-mail: changyuqing@ise.neu.edu.cn
(**CHANG Yu-Qing** Professor at the College of Information Science and Engineering, Northeastern University. Her research interest covers operating performance optimality assessment and fault diagnosis of complex system.)

王 敏 东北大学信息科学与工程学院硕士研究生. 主要研究方向为复杂工业过程运行状态评价与非优原因追溯.
E-mail: wangmin_neu@hotmail.com
(**WANG Min** Master student at the College of Information Science and Engineering, Northeastern University. Her research interest covers operating performance optimality assessment and non-optimal cause identification.)

蔡庆宏 国家电网辽宁省电力有限公司研究员. 主要研究方向为复杂工业过程运行状态评价与非优原因追溯.
E-mail: caiqinghong_0711@hotmail.com
(**CAI Qing-Hong** Researcher at Liaoning Electric Power Co. LTD., State Grid. His research interest covers operating performance optimality assessment and non-optimal cause identification.)