

基于因果关系的 CBR 模型用于转炉炼钢静态控制

王心哲, 韩敏*

(大连理工大学 电子与信息工程学院, 辽宁 大连 116024)

摘要: 转炉炼钢静态控制是动态控制的基础, 准确的静态控制对于改善动态控制质量、提高终点命中率具有重要意义. 采用案例推理方法建立了转炉炼钢静态控制模型, 计算了吹氧量和冷却剂加入量; 并将因果关系分析用于案例属性的确定, 增强了属性选择的有效性, 降低了模型计算的复杂度. 对一座 150 t 转炉的实际生产数据进行仿真, 结果表明, 因果关系分析能够有效地确定案例属性, 基于案例推理的转炉炼钢静态控制模型是切实可行的.

关键词: 转炉炼钢; 静态控制模型; 因果关系分析; 案例推理

中图分类号: TF724 **文献标志码:** A

0 引言

转炉炼钢是一个降碳升温的过程, 主要目的是在吹炼终点时得到目标碳含量和温度满足工艺要求的合格钢水. 通常, 对转炉炼钢生产的控制包括静态控制和动态控制. 静态控制模型是动态控制模型的基础, 其精度直接影响到动态控制的效果, 如果静态模型不精确, 仅依靠吹炼末期的动态控制来调整碳温, 不利于终点碳含量和温度的命中^[1~3]. 因此, 静态控制模型的研究至关重要.

目前已有的静态控制模型主要有 4 种: 理论模型、统计模型、增量模型和神经网络模型, 但由于炼钢过程反应的复杂性和随机性, 许多影响因素很难用准确的数学方程或统计方法进行描述, 且神经网络模型受限于网络结构和参数的调整, 已有模型的控制精度受到很大影响^[4~6].

案例推理(case-based reasoning, CBR)方法是人工智能发展较为成熟的一个分支, 它基于历史经验和知识进行推理, 已在故障诊断、自动纠错系统和决策支持等领域取得成功应用^[7~11], 但在复杂工业过程的建模与控制方面仍处于探索阶段. 其主要思想是根据当前问题的描述, 在案例库

中搜索与当前问题相似的案例, 再利用案例库中搜索到的相似案例的解, 构造当前问题的解.

本文使用案例推理方法建立转炉炼钢静态控制模型, 并将因果关系分析引入案例推理系统, 分析各变量对解空间的影响程度, 选择对解空间影响较大的变量作为案例的条件属性.

1 模型的建立

案例推理是一种类比推理方法, 主要有以下步骤: 案例描述、案例检索、案例重用、案例修正和案例保存. 首先建立对问题(案例)的描述, 根据案例描述, 在案例库中进行检索, 找到最相似的若干案例; 再利用搜索到的相似问题的解构造当前问题的解; 最后, 将得到的解输出给用户, 如果成功解决问题, 则将此案例存入案例库. 基于案例推理的转炉炼钢静态控制模型结构如图 1 所示.

1.1 案例描述

案例描述是案例推理过程的基础, 是案例检索和案例重用有效性的保证. 案例描述就是确定用于描述案例的各属性和结构的过程, 通常根据背景知识或经验确定. 本文将因果关系分析引入案例推理系统, 用于对案例属性的选择.

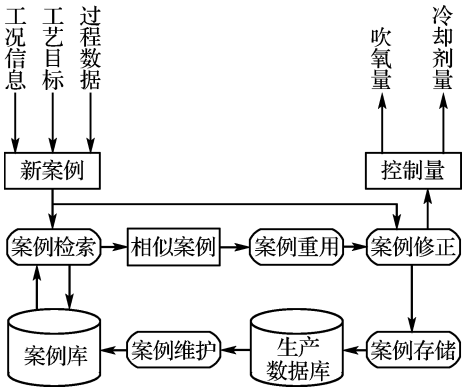


图 1 转炉炼钢案例推理静态控制模型结构

Fig. 1 Structure of CBR model for static control of converter steelmaking

1. 1. 1 因果关系 因果关系的思想最早由 Wiener 提出, 并由 Granger 将其公式化^[12], Granger 因果关系被广泛应用于各领域. 因果关系分析的主要思想是, 对于一个变量的当前值, 如果用该变量的过去值和第 2 个变量的过去值估计与仅由该变量自身的过去值估计相比, 得到的误差项的方差减小, 则认为第 2 个变量与该变量存在因果关系, 且第 2 个变量为该变量的因.

对于一个固定的变量 $x(t)$, 首先考虑基于 m 个历史测量值的自回归预测:

$$x(t) = \sum_{j=1}^m \alpha_j x(t-j) + \epsilon_x(t) \tag{1}$$

其中 $\epsilon_x(t)$ 为预测误差, 其大小可表示为方差形式 $\text{var}(\epsilon_x(t))$. 在自回归的基础上, 引入另一组固定变量 $y(t)$ 的历史测量值, 则可得到下面的预测:

$$x(t) = \sum_{j=1}^m \alpha_j x(t-j) + \sum_{j=1}^m \beta_j y(t-j) + \epsilon_{x|y}(t) \tag{2}$$

如果得到的预测误差的方差小于自回归预测误差的方差, 即 $\text{var}(\epsilon_{x|y}(t)) < \text{var}(\epsilon_x(t))$, 则变量 $y(t)$ 对变量 $x(t)$ 具有因果关系.

二变量间的因果关系分析虽然可以得到是否存在因果关系, 但无法确定是直接关系还是间接关系, 即存在如图 2 所示的两种情况:

其中, 图 2(a) 表示变量 z 对 x 的因果关系为间接关系, 是通过变量 y 实现的. 图 2(b) 则表示变量 z 与 x 间存在直接的因果关系. 为此, 使用条

件因果关系分析来确定变量间的因果关系属于哪种模式.

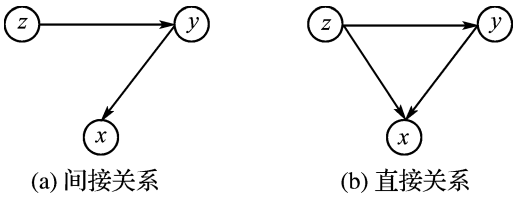


图 2 两种因果关系模式

Fig. 2 Two patterns of causal interactions

与二变量间因果关系的分析相似, 已知变量 z 与 x 之间存在因果关系, 在式(2) 的基础上引入变量 z 得到的预测表达式为

$$x(t) = \sum_{j=1}^m \alpha_j x(t-j) + \sum_{j=1}^m \beta_j y(t-j) + \sum_{j=1}^m \gamma_j z(t-j) + \epsilon_{x|yz}(t) \tag{3}$$

如果得到的预测误差的方差小于式(2) 的回归预测误差的方差, 即 $\text{var}(\epsilon_{x|yz}(t)) < \text{var}(\epsilon_{x|y}(t))$, 则变量 $z(t)$ 对变量 $x(t)$ 具有直接的因果关系, 否则说明具有间接的因果关系.

1. 1. 2 案例属性的确定 根据转炉炼钢生产的工况信息、工艺目标和过程数据构建案例的问题属性和解属性. 影响静态控制的因素很多, 如工艺目标要求、吹炼过程操作以及原料成分等. 转炉炼钢静态控制模型的案例描述包括两种属性:

(1)离散属性

离散属性主要考虑吹炼终点钢水的目标温度和碳含量. 根据操作规程, 冶炼的钢种不同时, 钢水终点目标温度和碳含量的高低有所不同. 因此, 案例的离散属性包括钢水的目标温度和目标碳含量. 由此, 离散属性集合可表示为{目标温度, 目标碳含量}.

(2)连续属性

连续属性的构建主要考虑设备信息和原料信息, 包括氧枪的枪龄、转炉的炉龄、铁水成分和温度、废钢装入量和辅原料加入量. 由此, 连续属性集合可表示为{设备信息, 铁水信息, 原料加入量信息}. 可用于描述连续属性的变量包括表示设备信息的氧枪的枪龄和转炉的炉龄; 表示铁水信息

的铁水的碳含量、硅含量、锰含量、磷含量、硫含量和铁水温度；表示原料加入量信息的废钢装入量、石灰加入量和白云石加入量。

如果选择所有的变量作为描述案例的属性，不仅会增加计算案例间相似度时的计算量，而且那些不重要的属性作为噪声会影响相似案例的选择，从而降低模型计算的精度。同时，案例属性过多还会使案例库的规模增长过快，不利于案例库的维护。由此，本文利用 Granger 因果关系和条件因果关系对案例的连续属性进行确定，选择重要的变量并剔除冗余变量，以提高案例推理系统的运行效率和有效性。

使用因果关系分析确定案例属性后，案例的构建描述为如图 3 所示的 4 层结构。

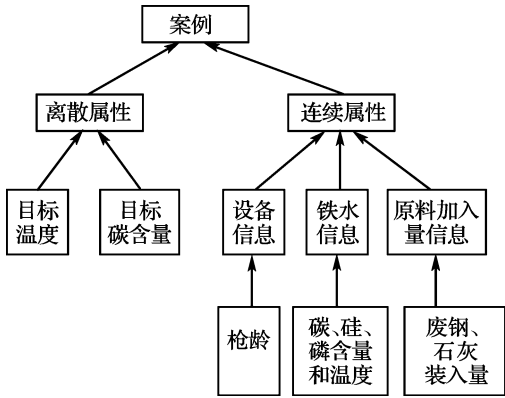


图 3 案例属性结构

Fig. 3 Structure of case attributes

1.2 案例检索

案例检索是案例推理过程的关键，是构造当前问题精确解的基础，其核心步骤是历史案例与当前案例间相似度的计算。目前较常用的案例检索方法有最邻近法、归纳索引法和知识引导法等^[13]。通过案例检索，获得与当前问题最相似的若干案例，检索得到的相似案例的解用于案例重用。

根据案例描述的结构，分别计算离散属性和连续属性相似度，整体相似度按照案例描述的 4 层结构计算。

1.2.1 离散属性相似度的计算 案例的离散属性与生产的工艺目标相关，主要包括钢水目标温

度和目标碳含量，这也是衡量钢水是否合格的两个指标。离散属性相似度的计算式如下：

$$sim_d = \frac{sim_{aim_c}(f, f_c) + sim_{aim_t}(f, f_c)}{2} \quad (4)$$

式中： sim_d 为离散属性的相似度； sim_{aim_c} 为目标碳含量的相似度， sim_{aim_t} 为目标温度的相似度。目标温度和目标碳含量相似度的计算式为

$$sim_{aim_C(t)}(f, f_c) = 1 - \frac{|f - f_c|}{\max\{f, f_c\}} \quad (5)$$

其中 f_c 为案例库中某个案例的离散属性值， f 为当前案例的离散属性值。

1.2.2 连续属性相似度的计算 与离散属性相似度的计算方法相同，连续属性相似度采用最邻近法计算，并且按照连续属性本身的 3 层结构进行计算。

顶层为连续属性相似度，相似度的计算表达式如下式所示：

$$sim_c = \sum_{i=1}^3 w_i sim_i \quad (6)$$

式中： sim_c 为案例连续属性的整体相似度； sim_i 为一个属性集合的相似度，分别表示设备信息、铁水信息、原料加入量信息这 3 个属性集合的相似度； w_i 为对应每个属性集合的相似度权值，由经验确定。

各属性集合中的属性元素均为连续型变量，采用如下的公式计算各属性集合的相似度：

$$sim_i = \frac{\sum_{j=1}^m (w_{ij} sim_{ij}(f_{ij}, f_{c,ij}))}{\sum_{j=1}^m w_{ij}} \quad (7)$$

式中： m 为第 i 个属性集合中属性的个数； f_{ij} 为当前炉次第 i 个属性集合中第 j 个属性值； $f_{c,ij}$ 为案例库中第 i 个属性集合中第 j 个属性值； $sim_{ij}(f_{ij}, f_{c,ij})$ 为第 i 个属性集合中第 j 个属性的相似度，即底层相似度，计算公式如式 (5) 所示； w_{ij} 为第 i 个属性集合中第 j 个属性相似度的权值，其值的确定过程如下。

首先，计算每个属性相似度的标准差：

$$\sigma_j = \sqrt{\frac{1}{n} \sum_{k=1}^n (sim_j^k - \overline{sim_j})^2} \quad (8)$$

其中 n 为案例库中案例的个数； sim_j^k 为当前案例

与案例库中第 k 个案例间第 j 个属性的相似度;
 $\overline{sim}_j$ 为当前案例与案例库中案例间第 j 个属性相似度的均值.

再利用式(8)的计算结果确定各属性权值的大小,如下式所示:

$$w_{ij} = \frac{\sigma_j}{\sum_{j=1}^n \sigma_j} \quad (9)$$

对于相同的属性,各个 sim_j^k 的值越接近,则 σ_j 越小,即不同的案例在该属性上的差异越小,也表示该属性在区分各个案例的相似程度上作用不大,则被赋予相对较小的权值.相反的情况被赋予较大的权值.

1.2.3 案例整体相似度的表达 案例整体相似度 sim_t 可由离散属性相似度和连续属性相似度的集合形式表达,即整体相似度包括离散属性相似度和连续属性相似度两个元素,表示为 $sim_t = \{sim_d, sim_c\}$.

对于当前案例,首先计算当前案例与案例库中案例间离散属性的相似度,选择满足一定离散属性相似度要求的案例;在此基础上,计算当前案例与选择案例间的连续属性相似度,并将计算结果按相似度的大小降序排序,选择与当前案例连续属性相似度值最大的 k 个历史案例作为最终的相似案例,相似案例的解在下一步的案例重用中使用.

根据案例描述的结构,将案例检索过程按照离散属性和连续属性分两步进行,基于离散属性的案例检索是连续属性案例检索的基础,案例检索的最终结果由连续属性的案例检索结果直接决定.

1.3 案例重用和修正

案例重用为当前问题提出建议解或者最终解,案例重用的方式直接影响到对当前问题解的精确性.

通过案例检索得到与当前问题最相似的 k 个案例的解,使用相似案例的解构造当前问题的解,计算表达式为

$$S = \frac{\sum_{i=1}^k (w_i s_i)}{\sum_{i=1}^k w_i} \quad (10)$$

式中: S 为构造的当前案例解; s_i 为第 i 个相似案例的解; k 为相似案例的个数; w_i 为权值,其大小为 sim_c .

1.4 案例存储和维护

案例的存储和维护对案例推理系统的运行十分必要.案例存储,可以丰富案例库中的案例.但随着生产的进行会造成案例库的过度膨胀,降低系统的运行效率,需要对案例库进行维护,以保证运行效率.如果当前炉次的吹炼结果满足工艺要求,且与历史案例的相似度小于一个阈值,则本炉次作为一个成功案例存入案例库.

2 仿真研究

使用某钢厂一座 150 t 转炉现场采集的实际生产数据进行仿真研究,选择 630 炉不使用复吹且终点碳含量和温度均命中的炉次.选择前 500 炉数据作为案例库中的案例,对其余 130 组进行静态控制计算.

使用 1.2 的案例检索方法在案例库中搜索与第 501 炉相似的炉次案例,选择与当前炉次最相似的前 13 个案例的解计算当前炉次的吹氧量和冷却剂加入量.计算结束后,将第 501 炉数据存储于案例库中,作为成功案例用于对下一炉的计算,依此类推,完成所有 130 炉的静态控制计算,并保证案例的更新.

2.1 模型计算结果

按照上述方法对实际数据进行仿真,得到静态吹氧量的模型计算值与实际值之间的均方根误差为 302.36 m^3 ,误差限度为 500 m^3 时的准确率为 90.00%;冷却剂加入量的模型计算值与实际值之间的均方根误差为 0.65 t ,误差限度为 1 t 时的准确率为 86.15%.使用案例推理计算的吹氧量 Q_o 和冷却剂加入量 Q_c 与实际值间的曲线图分别如图 4 和 5 所示.

2.2 与选择全部变量作为案例属性模型的比较

将因果关系分析用于对案例属性的确定,由此减小了案例描述的复杂度,且能控制案例库的规模,有利于案例库的维护.为了进一步验证该方法的有效性,将应用因果关系分析确定案例属性的模型(C-CBR)计算结果与选择全部变量作为案

例属性的模型(CBR)计算结果进行比较,结果如表1所示.

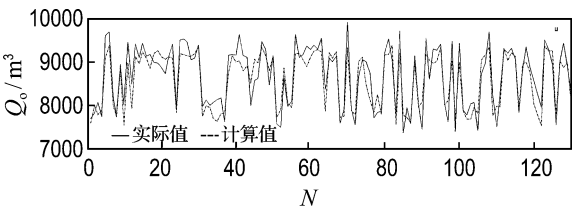


图4 吹氧量计算值与实际值间曲线

Fig. 4 Curves of calculation values and real values of oxygen blown amount

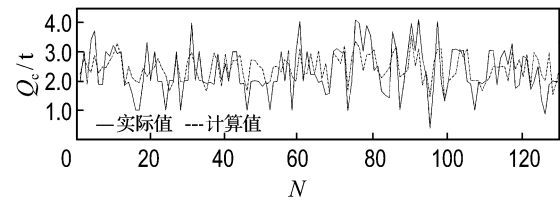


图5 冷却剂加入量计算值与实际值间曲线

Fig. 5 Curves of calculation values and real values of coolant addition amount

表1 不同案例描述时的模型计算结果比较

Tab. 1 Comparison of the results with different case description

模型	均方根误差		准确率/%	
	吹氧量/ m^3	冷却剂量/ t	吹氧量	冷却剂量
CBR	305.92	0.70	89.23	85.38
C-CBR	302.36	0.65	90.00	86.15

表1中,吹氧量的准确率的误差范围为小于 500 m^3 ,冷却剂量的准确率的误差范围为小于 1 t .

由表1可以看出,引入因果关系分析对案例属性确定后,不会影响模型的计算精度.

2.3 与现有方法的比较

将使用案例推理方法建立的静态控制模型的计算结果与文献[4、5]中的静态控制方法进行比较.文献[4]中对 120 t 转炉使用增量回归方法建模;文献[5]中对 250 t 转炉使用遗传算法结合径向基神经网络方法(GA-RBF)计算,比较结果分别见表2和3.

通过比较可以看出,使用案例推理方法建立

的转炉炼钢静态控制模型无论在精度还是在准确率上均优于已有方法.

表2 吹氧量的比较

Tab. 2 Comparison of oxygen blown amount

模型	均方根误差/ m^3	$<500\text{ m}^3$ 准确率/%
增量模型[4]	326.13	84.31
GA-RBF[5]	310.25	87.69
C-CBR	302.36	90.00

表3 冷却剂加入量的比较

Tab. 3 Comparison of coolant addition amount

模型	均方根误差/ t	$<1\text{ t}$ 准确率/%
增量模型[4]	0.82	78.46
GA-RBF[5]	0.75	83.08
C-CBR	0.65	86.15

3 结 论

使用因果分析方法对案例属性进行确定,能够在保证模型精度的前提下减小案例描述的复杂度,降低案例库的存储空间,提高案例库的维护效率.仿真结果显示,使用该方法建立的静态控制模型的效果优于已有方法,对实际生产具有一定的指导意义.

参考文献:

[1] 黄金侠,金宁德. 转炉冶炼终点静态控制预测模型[J]. 炼钢, 2006, **22**(1):45-48

[2] 陈忠伟,袁守谦. LD转炉冶炼的静态数学模型及实现[J]. 炼钢, 2000, **16**(5):31-34

[3] 朱光俊,梁本川. 转炉炼钢静态控制优化模型[J]. 炼钢, 1999, **15**(4):25-28

[4] 孟祥宁,张海鹰,朱苗勇. 转炉炼钢过程静态控制模型的改进[J]. 材料与冶金学报, 2004, **3**(4):246-249

[5] 王建辉,徐林,方晓柯,等. 基于一类GA-RBF神经网络的转炉炼钢静态模型控制[J]. 东南大学学报, 2005, **35**(2):90-94

[6] 丁容,刘浏. 转炉炼钢过程人工智能静态控制模

- 型[J]. 钢铁, 1997, **32**(1):22-26
- [7] CHANG Pei-chann, LIU Chen-hao, LAI R K. A fuzzy case-based reasoning model for sales forecasting in print circuit board industries [J]. **Expert Systems with Applications**, 2008, **34**(3):2049-2058
- [8] HUANG Mu-jung, CHEN Mu-yen, LEE Show-chin. Integrating data mining with case-based reasoning for chronic diseases prognosis and diagnosis [J]. **Expert Systems with Applications**, 2007, **32**(3): 856-867
- [9] NEGNY S, LE LANN J-M. Case-based reasoning for chemical engineering design [J]. **Chemical Engineering Research and Design**, 2008, **8**(6):648-658
- [10] GUILLERMO C R, NEGNY S, LE LANN J-M. Case-based reasoning and TRIZ: A coupling for innovative conception in chemical engineering [J]. **Chemical Engineering and Processing**, 2009, **4**(1): 239-249
- [11] TUNG Yuan-hsin, TSENG Shian-shyong, WENG Jui-feng, *et al.* A rule-based CBR approach for expert finding and problem diagnosis [J]. **Expert Systems with Applications**, 2010, **37**(3):2427-2438
- [12] GRANGER C W J. Investigation causal relations by econometric models and cross-spectral methods [J]. **Econometrica**, 1969, **37**(3):424-438
- [13] 李伟明, 穆志纯. 基于案例推理的冷轧负荷分配模型研究[J]. 系统仿真学报, 2006, **18**(8):2162-2165

Causality-based CBR model for static control of converter steelmaking

WANG Xin-zhe, HAN Min*

(School of Electronic and Information Engineering, Dalian University of Technology, Dalian 116024, China)

Abstract: Converter steelmaking static control is the basis for dynamic control. The accurate static control for the improvement of dynamic control and the improvement of endpoint hit percentage is of great significance. Case-based reasoning method is adopted for establishing the converter steelmaking static control model to calculate the oxygen blowing quantity and the addition of coolant. Causality analysis is used to establish the attributes of the case to enhance the validity of attributes selection and reduce the complexity of model calculation. Simulations are implemented using real production data of a 150 t converter and results show that case attributes selected by causality analysis are effective and case-based reasoning converter steelmaking static control model is feasible.

Key words: converter steelmaking; static control model; causality analysis; case-based reasoning