

Apriori 算法在分析反应器单元中的应用

1 问题背景

在实际的流程工业中，不可避免的存在一些非闭环的回路，在这些回路中，由于工况变化，使得期望的目标值偏移稳定值，造成系统不稳定或产品不合格，因此需要通过调整操作变量使目标值回到原来的稳定状态，以便得到合格的产品。随着工业现场总线技术的发展和计算机集散控制系统日渐成熟，使得大量工业数据可以通过成熟的数据采集和存储系统保存下来，而大量的操作数据中又隐含了大批的操作信息，可以利用这些信息，辅助工业生产。采用关联规则寻找操作变量与被控变量之间的关系，挖掘操作员应对现场工况的经验，用于调整生产、优化操作，对实际生产具有一定的指导意义。

在工业过程中，由于工业过程数据噪声大、易抖动的特点，其有效数据比较少，而 Apriori 算法简单、易理解、对数据要求低，擅长处理此类数据。该算法生成的规则可以根据预定义的最小支持度和最小置信度来控制规则的数目和质量，直观明朗，容易找到高质量的规则。

2 问题描述

选取某实际化工生产过程的反应器单元进行分析，如下图 1 所示。

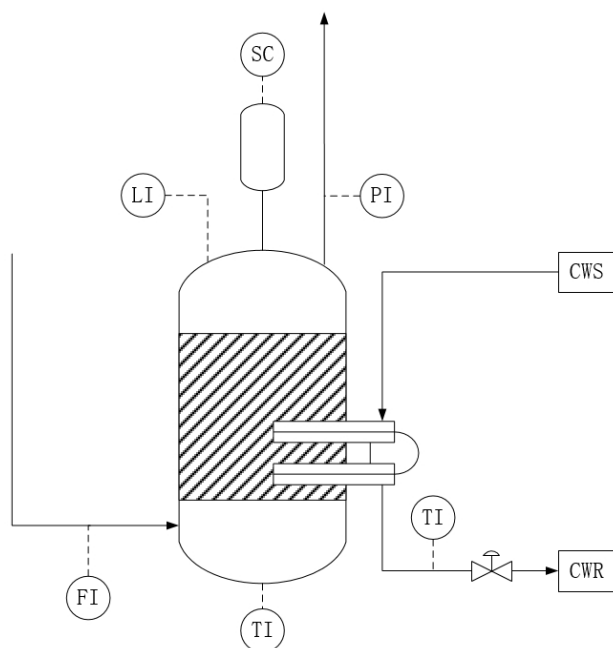


图 1 反应器单元

反应器内涉及 4 个化学反应：



其中 A、C、D、E 为气态物料参与反应，得到 G、H 两种产品，同时还生成副产物 F，整个反应还需要少量的惰性气体 B 的催化，所有反应均为放热反应。反应器的运行状态可通过压力、液位、温度三个状态变量反映，状态变量受反应器操作变量的影响。本案例分析选取六个操作变量（如下表 1），采用 apriori 算法探究其与反应器压力的关联关系。

表 1 操作变量

变量编号	过程变量名称	基准值	物理单位
XM1	D 进料量	63.03	kg ^h ⁻¹
XM2	E 进料量	54.01	kg ^h ⁻¹
XM3	A 进料量	24.72	kscmh
XM4	排放阀	40.08	%
XM5	反应器冷却水流量	41.09	m ³ h ⁻¹
XM6	冷凝水冷却水流量	18.22	m ³ h ⁻¹

3 方法和结果

(1) 数据预处理

首先对反应器单元每个变量选取 500 条采样数据，将反应器压力数值按照要求划分为低、中、高三个范围，并以 H1，H2，H3 标签进行区分。

表 2 压力数值标签

标签	H1	H2	H3
压力/kpa	P<2701.2	2701.2<P<2709.5	P>2709.5

由于反应器过程操作变量数据 XM1、XM2、XM3、XM4、XM5 以及 XM6 是在区间 [0, 100] 之间连续变化的实数，而关联规则算法只能为布尔或数值输入，因此采用 1σ 准则对各变量进行划分，得到各变量的分割点如表 3 所示。

表 3 各变量分割点数据

XM1	XM2	XM3	XM4	XM5	XM6
61.20	52.70	17.70	35.40	39.50	14.00
62.47	53.58	21.90	38.64	39.58	16.73
63.59	54.44	27.54	41.52	40.63	19.71
64.50	55.20	33.70	44.60	42.70	23.20

其中，XM11 表示 XM1 属于第一区间，即 $XM11 \in [61.20, 62.47]$ ，XM12 表示 $XM12 \in [62.47, 63.59]$ ，XM13 表示 $XM13 \in [63.59, 64.50]$ ，其余变量以此类推。

(2) 操作关联规则挖掘

以布尔化数据输入，并采用 Apriori 算法进行关联规则的挖掘。设置最小支持度 $\text{minsup}=0.01$ ， $\text{mincof}=0.7$ ，得到以下表 4、表 5 所示规则。

表 4 是关于反应器压力处于正常控制范围内的规则。

表 4 压力处于常值时的规则

	If	A	Then	B	Support	Confidence
1	If	XM11, XM31, XM62	Then	H2	1.198%	100%
2	If	XM23, XM33, XM41	Then	H2	1.397%	100%
3	If	XM22, XM31, XM41, XM61	Then	H2	1.197%	100%
4	If	XM33, XM61	Then	H2	2.196%	91.67%
5	If	XM11, XM23	Then	H2	1.796%	90.00%
6	If	XM11, XM23, XM41	Then	H2	1.796%	90.00%
7	If	XM33, XM41, XM61	Then	H2	1.796%	90.00%
8	If	XM11, XM31	Then	H2	1.597%	88.89%
9	If	XM22, XM33, XM61	Then	H2	1.597%	88.89%
10	If	XM33, XM41, XM53, XM61	Then	H2	1.597%	88.89%
11	If	XM11, XM31, XM41	Then	H2	1.397%	87.50%
12	If	XM11, XM23, XM32, XM41	Then	H2	1.397%	87.50%
13	If	XM11, XM23, XM53	Then	H2	1.198%	85.71%
14	If	XM11, XM33, XM53	Then	H2	1.198%	85.71%

表 5 是关于反应器压力处于目标值以上的规则。

表 5 压力高于常值时的规则

	If	A	Then	B	Support	Confidence
1	If	XM13, XM22, XM31, XM41, XM62	Then	H3	1.198%	75.00%

(3) 规则分析

将计算所得的操作规则信息以该变量所属区间的形式表示出来，即规则反布尔化，表 6、表 7 分别为上述规则反布尔化结果。

表 6 是反应器压力处于正常范围内部分规则的反布尔化结果。

表 6 压力处于常值时的反布尔化规则

	A	B	Support	Confidence
1	XM1 ∈ [61.20, 62.47], XM3 ∈ [17.7, 21.90] XM6 ∈ [16.73, 19.71]	H2	1.198%	100%
2	XM2 ∈ [54.44, 55.20], XM3 ∈ [27.54, 33.70] XM4 ∈ [35.40, 38.64]	H2	1.397%	100%
3	XM3 ∈ [27.54, 33.70], XM6 ∈ [14.0, 16.73]	H2	2.196%	91.67%
4	XM1 ∈ [61.20, 62.47], XM2 ∈ [54.44, 55.20]	H2	1.796%	90.00%
5	XM1 ∈ [61.20, 62.47], XM2 ∈ [54.44, 55.20] XM4 ∈ [35.40, 38.64]	H2	1.796%	90.00%
6	XM3 ∈ [27.54, 33.70], XM4 ∈ [35.40, 38.64] XM6 ∈ [14.0, 16.73]	H2	1.796%	90.00%
7	XM1 ∈ [61.20, 62.47], XM3 ∈ [17.7, 21.90]	H2	1.597%	88.89%
8	XM2 ∈ [53.58, 54.44], XM3 ∈ [27.54, 33.70] XM6 ∈ [14.0, 16.73]	H2	1.597%	88.89%
9	XM1 ∈ [61.20, 62.47], XM3 ∈ [17.7, 21.90] XM4 ∈ [35.40, 38.64]	H2	1.397%	87.50%
10	XM1 ∈ [61.20, 62.47], XM2 ∈ [54.44, 55.20] XM5 ∈ [40.63, 42.70]	H2	1.198%	85.71%
11	XM1 ∈ [61.20, 62.47], XM3 ∈ [27.54, 33.70] XM5 ∈ [40.63, 42.70]	H2	1.198%	85.71%

表 7 是反应器压力处于目标值以上规则的反布尔化结果。

表 7 压力高于常值时的反布尔化规则

	A	B	Support	Confidence
	XM1 ∈ [63.59,64.50],XM2 ∈ [53.58,54.44]			
1	XM3 ∈ [17.70,21.90],XM4 ∈ [35.40,38.64]	H2	1.198%	75.00%
	XM6 ∈ [16.73,19.71]			

由上述规则反布尔化结果可知，要将反应器压力控制在一个正常的范围内，需要使原料 D 和原料 A 的进料量同时处于一个较低的水平，若原料 A 的进料量增加，则需要增大反应器冷却水的流量，降低反应器内温度，促使放热反应进行，从而降低反应器内压力。当反应器内压力高于正常范围时，需要降低原料 A 的进料量，增加原料 E 的进料量，促使气体原料转为液体产物，同时增加冷凝水用量，降低反应器内温度，最终可使压力回归正常值。

综上，通过 Apriori 算法挖掘到的反应器内状态变量与操作变量之间的关联关系与化工过程实际情况相符，得到的结果与工艺理论一致，因此关联规则可用于指导生产，具有实际应用价值。