

汽车涂装车间质量缺陷关联规则及预测方法研究

刘 博, 林 涛, 吕朋辉, 张程皓, 梁 胜

(中国汽车工业工程有限公司, 天津 300113)

摘要: 涂装车间喷涂质量缺陷是良品率难以突破的瓶颈因素。车身质量缺陷受制于复杂的工艺、设备以及物料和环境等因素, 不同类型的缺陷发生的概率难以预测。传统六西格玛与 SPC 方法侧重单变量监控, 难以揭示“多事件耦合触发缺陷”的复杂机制。本文提出一种基于 FP-Growth 的缺陷-事件关联挖掘框架: 首先将工艺报警、设备事件、新批次物料更换事件与质量缺陷进行“事件化”与“时间窗拼接”, 构建交易型数据集; 随后利用 FP-Growth 高效提取高频缺陷前兆规则; 最后设计质量缺陷事件预测服务, 实现缺陷提前 5~10 min 的发生概率推测并预警。该研究为涂装车间质量管理智能化转型提供了可复制的轻量级算法范式。

关键词: 涂装缺陷; FP-Growth; 关联规则; 实时预警; 数据挖掘

中图分类号: TQ639

文献标志码: A

文章编号: 1007-9548(2026)02-0051-05

Research on Association Rules and Prediction Methods for Quality Defects in Automotive Painting Workshops

LIU Bo, LIN Tao, LYU Peng-hui, ZHANG Cheng-hao, LIANG Sheng

(Automotive Engineering Corporation, Tianjin 300113, China)

Abstract: Painting workshops coating defects are the key bottleneck that keeps first-pass yield from improving. Car-body quality defects are governed by a complex interplay of process parameters, equipment states, material changes and environmental factors, making the probability of any specific defect hard to forecast. Traditional six-sigma and SPC methods focus on single-variable monitoring and fail to uncover the mechanism in which "multiple coupled events trigger a defect". We propose an FP-Growth-based framework that mines defect-event associations. First, process alarms, equipment events and new-batch material-change events are "eventized" and "time-window-aligned" with quality defects to create a transactional data set. FP-Growth is then used to extract high-frequency precursor rules efficiently. Finally, a defect-event prediction service is designed that estimates the probability of a defect 5~10 minutes ahead and issues an early warning. The study offers a lightweight, reusable algorithmic paradigm for the intelligent transformation of painting workshop quality management.

Key words: coating defects; FP-Growth; association rules; real-time early warning; data mining

0 引言

在汽车厂涂装车间, 如何精准预测喷涂缺陷并实施闭环控制, 始终是喷涂质量管理长期研究的课题。然

而, 涂装工艺非常复杂, 过程涉及前处理、电泳、电泳烘干、密封胶、中涂、中涂烘干、色漆、清漆、面漆烘干等十余道工序, 车身按照工艺顺序在工艺段之间穿梭。涂装线机器人数量众多, 相关工艺参数呈多维、海量的特征。人、机、料、法、环任意一个环节的变化点均可能在后续工位放大, 形成颗粒、缩孔、流挂、色差等缺陷。其中多半的缺陷源于“多事件耦合”, 传统 SPC 控制图仅能对单变量超限进行报警, 无法刻画“多事件耦合”与

收稿日期: 2025-12-22

作者简介: 刘博(1994—), 男, 本科, 工程师, 主要从事汽车行业涂装车间智能应用软件的开发工作。E-mail: 18843103252@163.com。

缺陷之间的非线性关系。尽管车间内部报警机制完善,可覆盖绝大多数生产场景,却因与质量缺陷关联度不足而丧失预测价值。因此,亟需一种数据驱动、可解释、低延迟的方法,从海量历史数据中提取“事件组合→缺陷”的强规则,实现前置干预。这使涂装车间的质量管理从“被动处理”转向“主动预防”新模式。

1 相关工作

关联规则 (Association Rule) 概念由 Agrawal 于 1994 年在“购物篮分析”中首次提出^[1],旨在从无序交易中发现“啤酒→尿布”式共现规律。关联规则挖掘则是以某种方式分析数据源,从数据源中发现关联或相关关系,即从数据源中发现高频共现的项目集,也称频繁项集(简称频繁集),然后再利用这些频繁集产生关联规则^[2]。随后,Apriori、FP-Growth 等高效算法相继出现,使其迅速渗透到零售、金融、电信、医疗乃至工业制造等海量事务型场景。与以“标签-损失”为核心的机器学习不同,关联规则无需预先标注样本,仅通过支持度-置信度框架即可输出可解释、可审计的“IF-THEN”知识,因而天然适合缺陷根因分析、故障先兆捕捉等工业任务。目前关联规则算法在半导体、电力、注塑等多个领域已验证其挖掘“缺陷先兆”的有效性。刘智等^[3]在智能电表质检业务中利用关联规则

已有应用,李济同^[4]在船舶涂装领域的防腐效果检测中结合关联规则算法也有相关研究,但在汽车涂装领域的应用仍面临四大挑战。

1)多源异构数据的获取:涂装工艺复杂,生产过程产生大量数据。例如 PLC 报警、机器人事件、新批次物料事件、人工保洁事件等,这些数据呈现数据来源不同且数据结构存在巨大差异的特性。

2)样本不平衡:特殊缺陷占比<1%,易被视为噪声。

3)高频计算:按照节拍 60 JPH 计算,算法需在分钟级完成规则匹配。

4)事件滞后性:从导致缺陷发生的事件到缺陷受工艺影响存在较大的时差,具有明显的滞后性。

常用的关联规则算法包括 Apriori 算法和 FP-Growth 算法,FP-Growth 算法^[5]是对 Apriori 算法最经典的改进。涂装车间的各类数据规模大,且数据种类丰富。因为 FP-Growth 算法无法有效挖掘连续性数据,对于离散型数据更能准确地挖掘出数据之间的关联规则^[6]。因此需要针对涂装车间内各类数据进行事件化预处理。本文以 FP-Growth 为核心,针对上述挑战提出系统化解方案。结合质量管理体系提出“离线关联挖掘+在线规则匹配”的系统架构,如图 1 所示。

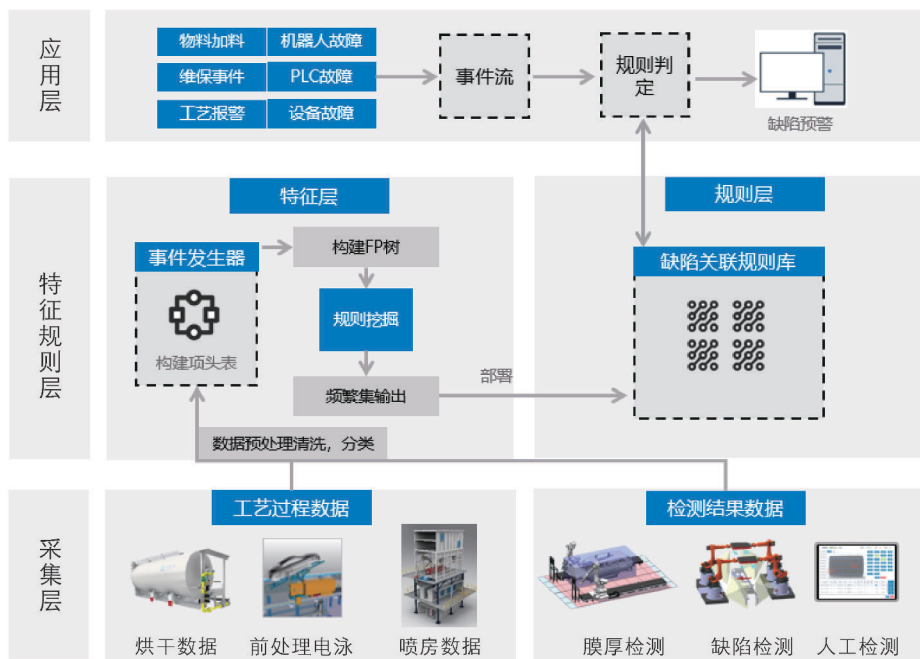


图 1 “离线关联挖掘 + 在线规则匹配”系统架构

1)数据层:工艺过程数据+检测结果数据,例如:工艺设备报警、物料加料记录、机器人报警、维保记录、车身缺陷记录等;

2)特征层:事件发生器+时间窗拼接,交易型数据集;

3)规则层:FP-Growth 离线训练,缺陷关联规则

库;

4)应用层:事件实时流、规则匹配、概率预警推送。

2 数据准备与事件化模型

频繁项集生成是 FP-Growth 算法生产关联规则挖掘过程中的关键步骤^[7]。而数据预处理与事件化生成是其中的关键。涂装生产过程中会产生大量的过程数据,例如各类工艺参数异常、设备故障、物料变更、人员变化等。这类数据与车身质量缺陷存在千丝万缕的联系。因此为实现关联规则输出,我们必须进行前序工艺的事件化处理,以及与各类缺陷发生时间进行窗口滑动匹配,最后形成简单的因果关系对应。首先对连续变量进行分箱离散化,将温度、湿度、压力等工艺参数转化为“超上限”“正常”“低于下限”等事件标签;随后把机器人报警、物料批次切换、保洁作业等异构日志统一编码为布尔型事务,实现秒级特征压缩,既保留工艺物理意义,又满足 FP-Growth 对离散项的输入要求,为后续高置信度规则挖掘奠定基础。

2.1 报警事件与缺陷分类

2.1.1 缺陷分类

涂装工艺缺陷种类繁多,且各自成因不同。图 2 为某新能源车企涂装车间的各类缺陷占比,其中钣金、凹坑、打磨印、划伤等缺陷,受来料、人工操作等影响较大,且负样本较少,存在较大的不确定性,因此这类缺陷不参与关联规则训练。

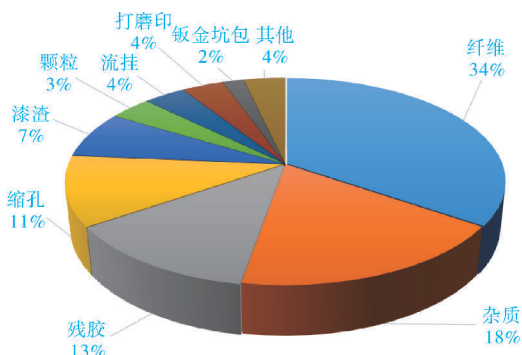


图 2 车身缺陷类型占比

另外,参考吴文娇^[8]针对汽车涂装漆膜缺陷的分析,结合某新能源车企涂装工艺专家给出的建议,挑选出高频出现且与前序工艺存在明显因果关系的缺陷作为数据挖掘样本。最终选取杂质、缩孔、流挂、色差、气泡、橘皮这六类缺陷,其触发事件在 MES 中记录完整,时间窗对齐后样本充足,可满足 FP-Growth 对支持度的最低要求,并避免小样本带来的假关联风险,如表 1 所列。

表 1 涂装车间常见车身缺陷类型

类型	编码	中文描述
缺陷	D_1	缩孔
缺陷	D_2	橘皮
缺陷	D_3	流挂
缺陷	D_4	色差
缺陷	D_5	杂质
缺陷	D_6	气泡

2.1.2 报警事件及变化点

按涂装工艺专家的意见提炼车间内重要报警及变化点(见表 2),作为关联规则数据挖掘的子项事件输入(该事件作为试验数据,选取车间部分报警及变化点,不代表真实车间的全部事件)。

表 2 涂装车间高等级报警事件及变化点

类型	编码	中文描述
报警	A_1	烘房输送故障
报警	A_2	喷涂流量异常
报警	A_3	喷房温湿度异常
报警	A_4	机器人故障
报警	A_5	喷涂压力异常
报警	A_6	流平时间不足
报警	A_7	烘干温度异常
事件	E_1	新批次物料更换
事件	E_2	喷房室体保洁
事件	E_3	空调过滤袋更换
事件	E_4	油漆黏度异常
事件	E_5	机器人旋杯更换

2.2 事件化规则

本文所讨论的预测方法属于质量缺陷的概率预测。实际生产过程中即便没有变化点及报警发生在车身不同部位也有概率造成缺陷。因此我们将小概率发生缺陷以及偶发性缺陷视作噪声处理。除此之外,涂装车间不少岗位仍然是完全由手工完成的,如电泳打磨、中涂打磨、PVC 的粗细密封等^[9]。这类事件难以量化,不好做事件化处理,因此不参与 FP 树的构建。作为试验数据文中所用到的报警事件判定规则已隐藏企业敏感阈值。

3 FP-Growth 关联挖掘

3.1 算法执行

本文采用 FP-Growth 算法挖掘“涂装质量缺陷等价事件”之间的关联规则,核心思想是:先利用 FP-Tree 压缩交易库,再递归提取频繁项集,最后转换为缺

陷预警规则。具体步骤如下所示。

设对于任意频繁项集 $Z \in F$, F 为涂装车间内全体事件项的集合。若存在缺陷项 $D_j \in Z$, 则拆分规则为 $X=Z\{D_j\} \rightarrow Y=\{D_j\}$, 其中 X 称为前件(报警事件/变化点信号), Y 称为后件(质量缺陷信号), 规则含义为: 当报警事件集 X 出现时, 导致质量缺陷 Y 共现。

$\mathcal{J}=\{A_1, \dots, A_7, E_1, \dots, E_5, D_1, \dots, D_6\}$ 为报警事件, 车间变化点和导致其发生的质量缺陷编码集合;

$\mathcal{D}=\{T_1, \dots, T_N\}$ 为滑动窗交易库, 具体时间滑动窗口由涂装工艺决定。本次试验样本 $|\mathcal{D}|=1\ 000$;

$W(D_j)=1/P(D_j)$ 为缺陷权重, 其中 $P(D_j)$ 为缺陷 D_j 在全体样本中发生的概率。例如缩孔 D_1 在单位样本中发生的概率为 1%, 则它的缺陷权重约 10^2 , 保证稀有项进入频繁集;

\min_sup 最小加权支持度=1%, \min_conf 最小值置信度=80%, $Kulc_th$ 修建阈值=0.7, 以上数据基于行业经验以及工艺专家审定。

1) 加权支持度(共现概率)

$$sup_w(Z) = \frac{\sum_{i=1}^{|\mathcal{D}|} I(Z \subseteq T_i) \cdot w(T_i)}{\sum_{i=1}^{|\mathcal{D}|} w(T_i)}, w(T_i) = \sum_{v \in T_i \cap D_1, \dots, D_6} w(v)$$

\mathcal{D} 为测试数据集中的 1 000 条 5 min 窗口交易库, T_i 为第 i 个窗口内的事件列表(例如 $\{A_3, E_1, D_1\}$), $I(Z \subseteq T_i)$ 为指示函数, 如果窗口 T_i 完全包含项集 Z , 取 1, 否则取 0。

加权支持度=含 Z 的窗口权重和÷总权重和, 就是 Z 的加权共现频率。

$w(T_i)$ 只对窗口里的缺陷项求权重, 报警/事件不计权重。

$w(v)=1/S$, 其中 S 为该缺陷在全体窗口中的出现比例。

例如: 缩孔 D_1 出现 240 次 $\rightarrow w(D_1) = 240\ 000/240 = 1\ 000$ 。因此“含缩孔”的窗口权重瞬间变大, 保证缩孔能进入频繁项集。

2) FP-Tree 节点累加(FP-Growth 核心)

$insert(T_i, root): root.count += w(T_i)$, 对 T_i 按 sup_w 降序递归插入子节点。

3) 频繁项集提取(FP-Tree 递归输出)

$F=Z \subseteq \mathcal{J}; |sup_w(z)| \geq \min_sup, Z \subseteq \mathcal{J}$ 表示任意子集(例如 $\{A_3, E_1, D_1\}$ 或 $\{A_5, D_3\}$)。只要子集 Z 的加权频率 $\geq 1\%$, 就认定它是频繁项集。

4) 规则生成(拆分缺陷后件)

对任意 $Z \in \mathcal{J}$ 且 $D_j \in Z$, 构造 $X=Z\{D_j\}, Y=D_j$, 规则 $X \Rightarrow Y$ 。

5) 支持度 & 置信度(直接复用 FP-Tree 计数)

$$支持度\ sup(X \Rightarrow Y) = \frac{count_{Tree}(X \cup Y)}{|\mathcal{D}|}$$

$$置信度\ conf(X \Rightarrow Y) = \frac{count_{Tree}(X \cup Y)}{count_{Tree}(X)}$$

$count_{Tree}(\cdot)$ 为 FP-Tree 节点累加值, 无需再扫描数据库。

6) Kulc 系数(冗余修剪)

FP-Growth 关联规则挖掘过程中会有大量的关联规则生成, 但实际上有很多无效甚至错误规则输出, 这需要进一步筛选有效判定规则。

$$Kulc(X \Rightarrow Y) = \frac{conf(X \Rightarrow Y) + conf(Y \Rightarrow X)}{2}, conf(Y \Rightarrow X)$$

称为反置信度。若 $Kulc < 0.7$, 则判定为弱关联, 予以剔除; $Kulc \geq 0.8$ 保留并进入缺陷预警关联规则库。

表 3 给出了经过 $Kulc$ 修剪过后 Top-10 高置信度规则示例。

表 3 Top-10 高置信度缺陷与事件关联规则

规则 ID	前件(报警/事件)	后件(缺陷)	置信度/%	支持度/%
R01	{A ₁ , A ₃ , E ₁ }	D ₁ (缩孔)	82	68
R02	{E ₂ , E ₃ }	D ₅ (杂质)	79	71
R03	{A ₄ , E ₄ }	D ₆ (气泡)	77	80
R04	{A ₇ , E ₃ , E ₅ }	D ₅ (流挂)	72	82
R05	{A ₁ , E ₁ }	D ₄ (色差)	76	93
R06	{A ₂ , A ₃ , A ₄ }	D ₂ (橘皮)	78	90
R07	{E ₁ , E ₂ }	D ₅ (杂质)	88	73
R08	{A ₅ , A ₆ }	D ₃ (流挂)	81	70
R09	{A ₄ , A ₃ }	D ₆ (气泡)	70	71
R010	{A ₅ , E ₁ , E ₄ }	D ₁ (缩孔)	80	82

3.2 规则可解释性验证及结果过滤

邀请工艺专家对 R01-R10 进行因果复核。例如: 当喷涂机器人故障, 压力异常, 以及黏度问题时, 漆层太厚或太湿导致的漆膜从上向下流或下边缘增厚的现象, 大概率造成流挂, 与现场经验一致, 验证了规则的可信度。最终算法生成的规则将作为可解释的“IF-THEN”知识输出给车间质量管控系统, 进行下一步质量缺陷发生的概率预测。为进一步提升规则可用性, 系统同步计算支持度、置信度、提升度三大指标, 并设置置信度 $\geq 80\%$ 、提升度 ≥ 1.5 的硬阈值, 自动过滤伪相关; 同时对含“机器人报警+环境湿度”这类高共识项的规则标记“专家确认”标签, 确保推送至产线的每一条知识均可追溯、可复现、可审计。

4 质量预警

4.1 预警流程

实时报警/事件流经 5 min 滑动窗口聚合后,由质量管理系统在 1 s 内完成 FP-Tree 规则匹配(高 $Kulc \geq 0.7$ 和 $conf \geq 80\%$),命中即通过 WebSocket 推送缺陷概率与措施至精修工段长及质量工程师;现场扫码执行并回写处置结果到 MES,系统每日拉取新标签重新跑加权 FP-Growth,热替换规则库,实现“秒级预警-分钟处置-日级自迭代”闭环。

4.2 业务闭环

缺陷预警触发→精修工段长及质量工程师收到缺陷概率及推荐措施(如更换过滤袋、控制精修出车颜色比例)→MES 记录处置结果→用于下一轮规则评估,形成自迭代闭环。

最终,在线阶段将实时采集的“报警事件流”与规则库匹配,若命中任一高 $Kulc$ 规则,即输出缺陷类别及概率,实现事件→车身涂装质量缺陷发生概率秒级判定。

5 结论与展望

数据驱动且可解释的关联规则方法在复杂离散制造场景中存在非常高的落地价值。其中涂装车间的缺陷预测是非常典型的应用场景,除文本论述内容外,仍有以下拓展方向:

1)提出的“事件化+FP-Growth+实时匹配”框架可有效揭示涂装多源报警事件与质量缺陷的耦合机制,多源报警与缺陷的耦合机制被首次转译为“秒级可读的 IF-THEN 知识”,为后续车间专属知识库的构建提供重要参考;

2)与重型黑箱模型相比,规则具备可解释、可审计、可快速迭代优势,更契合汽车行业严苛的 FMEA 流程^[10],可将该功能纳入涂装车间日常运营管理体系中,进行考核;

3)未来将进一步引入序列规则(PrefixSpan)以刻画事件时序效应,并探索边缘计算部署的可能性,降低对云端服务的依赖。

本文研究的“事件化+FP-Growth+实时匹配”的涂装车间质量缺陷预测方法可应用于涂装质量管理系统,可提前给出缺陷概率提示;未来搭配机器人自动缺陷检测,有望提高一次检出率,缩短精修报交停线时长,为涂装车间质量管理提供强有力的数据支撑。

参考文献:

- [1] AGRAWAL R, SRIKANT R. Fast algorithms for mining association rules[C]. Proceedings of the VLDB Endowment, Santiago: Morgan Kaufmann Publishers, 1994: 487-499.
- [2] 王朝霞. 数据挖掘[M]. 北京: 电子工业出版社, 2018.
- [3] 刘智, 郝晓磊, 王娟, 等. 关联规则算法在智能电表质检业务的应用研究[J]. 工业控制计算机, 2025(9): 46-47.
- [4] 李济同. 基于关联规则算法的船舶涂装防腐效果检测方法[J]. 中国水运, 2024(21): 65-67.
- [5] AGRAWAL R, SHAFER J C. Parallel mining of association rules[J]. Journal of Supercomputing, 2007(3): 273-299.
- [6] 张斌, 滕俊杰, 满毅. 改进的并行 fp-growth 算法在工业设备故障诊断中的应用研究[J]. 计算机科学, 2018(S1): 508-512.
- [7] 杜世举, 郝晓娜, 杜阳, 等. 基于 FP-Growth 算法的 MES 系统生产关联规则挖掘技术研究[J]. 电脑编程技巧与维护, 2024(9): 30-32.
- [8] 吴文娇. 汽车涂装漆膜缺陷分析[J]. 现代涂料与涂装, 2024(1): 60-63.
- [9] 陈慕祖. 汽车涂装过程中的质量保证[J]. 现代涂料与涂装, 2011(12): 41-43.
- [10] 王楚悦. FEMA 工具在汽车涂装工艺中的应用[J]. 现代涂料与涂装, 2023(2): 69-72. ◆

2026 年《现代涂料与涂装》征订启事

《现代涂料与涂装》期刊是由中昊北方涂料工业研究设计院有限公司主办的全国性科技期刊,国内外公开发行,国际连续出版物号:ISSN 1007-9548,国内统一连续出版物号:CN 62-1135/TQ;本刊是中国学术期刊综合评价数据库来源期刊;《中国学术期刊(光盘版)》《中国期刊网》《万方数字化期刊群》《维普资讯》全文收录期刊;美国《化学文摘》(CA)收录期刊。突出实用性与理论性相结合的报道理念,侧重于企事业单位的研究成果传播,为实际生产遇到的问题提供参考和解决方案。本刊为月刊,每月 20 日出版,大 16 开本,彩版印刷,每期定价 15.00 元。

请根据您的方便,选择以下方式订阅:

1. 通过当地邮局订阅,国内邮发代号 54-65,全年 180 元。
2. 直接向本刊编辑部订阅,纸质版全年 282 元(含邮费),电子版全年 120 元。

汇款信息:

单位名称:中昊北方涂料工业研究设计院有限公司
 开户行:中国建设银行股份有限公司兰州拱星墩支行
 账号:6200 1360 0190 5150 0638



编辑部订阅二维码