

基于多维数据的设备画像构建及应用研究

何 育

国网江苏省电力有限公司设备部 江苏 南京 210000

摘 要: 设备画像是通过各种方式收集信息,将已有信息进行分析和挖掘,进而描述设备特征的一种方法。本文通过总结设备画像的相关概念和标签技术的相关知识和研究,利用三级标签的技术体系,综合考虑数据采集层,标签库层和标签应用层三个层次的架构体系,综合多维数据,详细分析资产全生命周期的数据情况,并注重标签的生成、管理、优化和分析,形成独特的资产全生命周期的研究方法,最终通过一系列综合评分方法精准刻画设备资产全貌,并实现基于设备画像的场景应用构建。

关键词: 设备画像; 标签; 场景应用构建

引言

在信息技术和大数据分析技术快速发展的今天,电网企业管理面临新的挑战。随着数据量的爆炸性增长,企业需要采用更高效的数据管理方法来应对全周期资产管理,这包括规划、设计、采购、建设、运维、财务和绩效评估等多个环节,需要更优秀的数据收集和管理方法来处理,其中关键在于如何整合不同系统的数据,构建标签体系,实现数据的可视化和应用。这不仅有助于深入理解资产状态,而且对于推动企业的持续健康发展至关重要。通过科学的数据整合和分析,企业能够制定更精准的管理策略,以应对日益增长的数据挑战^[1]。

针对目前面对的种种问题,基于资产管理数据底座数据贯通共享能力,本文利用大数据分析挖掘技术,以“标签”为形式,多视角、立体化的资产全生命周期设备画像,实现对设备特征的全面刻画。围绕安全、效能、成本三个维度,建立SEC单体设备综合评分指标体系,准确把握设备运行综合状态,实现资产全生命周期的精细化、差异化管理。最后,形成完整的关于全寿命资产周期管理的资产画像和标签体系,并形成一套完备的评分体系,以全面提升资产全生命周期管理数字化管理水平。

1 资产全生命周期数据全息模型的构建

结合数据采集接入过程,可构建基于设备画像的资产全生命周期数据全息模型,以实物“ID”为抓手,实现设备资产全方位关联、全流程贯通、全业务应用的数据资源共享生态,形成资产全生命周期各专业业务数据“同源聚合”、“同源共享”的数据融合新局面。

在构建过程中,结合资产管理业务协同需求,融合ERP、PMS、ECP、网上电网物资供应链、财务多维精益

等系统数据,梳理完善资产全生命周期关键信息数据范围,基于数据中台、资产管理数据底座,统一资产全生命周期各阶段、各专业数据物理模型、数据字典;建立数据标签及数据对应关系表,完成数据关系分析,连接跨专业数据断点,实现全量数据接入,确保数据全面贯通共享^[2]。

2 设备画像的构建与应用

设备画像是通过各种方式收集信息,将已有信息进行分析和挖掘,进而描述设备特征的一种方法。利用设备画像,有助于我们更好地描述资产状况、构建应用场景^[3]。

2.1 利用设备画像精准刻画设备资产全貌

基于资产管理数据底座数据贯通共享能力,利用大数据分析挖掘技术,以“标签”的形式,逐步构建设备层、站线层、组织层三个层次,形成立体化的资产全生命周期设备画像。通过对设备赋予不同安全、质量、成本等维度数据,加强资产管理业务数据分析与验证,实现对设备特征的全面刻画;围绕安全、效能、成本三个维度,建立SEC单体设备综合评分指标体系,支撑设备综合评价管理,准确把握设备运行综合状态,实现资产全生命周期的精细化、差异化管理,辅助项目精准投资决策、设备健康管理、全过程质量管控等场景应用。



图1 个体设备画像

通讯作者: 何育, 通讯邮箱: yvav@foxmail.com

2.2 基于设备画像的场景应用构建

1) 基于设备画像的变压器精细化管控策略应用。精细化运维策略是指建立在变压器画像标签体系的基础上，通过将变压器各类显性及隐形的特征进行描述组合，形成变压器的健康状态的画像，同时根据变压器的所属标签，结合国家电网的检修规程、业务专家经验制定应对的运维策略。

2) 基于设备画像的变压器运行状态预警。根据设备画像标签体系内的健康状况标签，通过大数据算法，可将设备健康水平分为三类，即健康设备、亚健康设备、病态设备，将不同状态的设备设置对应的运维策略，辅助运检决策管理。病态画像警示着变压器可能正处于一个不稳定阶段，其影响可能非常广泛，因此必须引起高度关注。亚健康画像则表明变压器虽然目前尚未达到危险状态，但其性能可能正在波动，存在向更严重转变的可能性。正常画像则表示变压器目前运行稳定，处于一个健康的运行周期。

3) 基于标签组合制定的综合风险评价系统。风险评估方法通常基于特定的标签组合来确定变压器的潜在风险。变压器的风险因素可以通过两种主要的标签来识别：一种是与风险发生可能性相关的标签，另一种是与风险影响范围相关的标签。利用标签规则将风险发生概率的运行年限划分成初期投入期、成熟稳定期及老化衰退期，对应的风险发生概率为低、中、高；将风险影响范围的电压等级划分成三个等级，对应的风险影响范围分为小、中、大^[4]。通过将两者进行组合，即可得出对变压器的风险评价结果。

3 标签体系的架构与分类

3.1 标签定义

画像的标签是一个同时包含研究主体的特定属性和具体数值（或数值区间）的二元组。画像标签与特征的不同之处在于画像标签还同时体现了属性。标签的属性包括：标签名称、ID、数据类型、权重、所属对象、使用周期、标签类型、创建日期、创建人、最近修改日期等等。

3.2 技术架构

设备画像的技术架构以单体设备为研究主体，共分为三层，分别是“数据采集层”“标签库层”和“标签应用层”。技术架构以数据汇集为基础，以标签生成、优化、分析及管理为核心，以标签应用为目标，利用大数据深度挖掘算法，整合设备全生命周期各阶段属性数据，开展设备画像标签体系构建^[5]。



图2 总体架构图

1) 数据层。作为构建设备画像标签体系的基石，数据采集层通过整合资产管理数据，实现信息的共享与流通。它收集了设备全生命周期的关键信息，这些数据为标签的生成和持续改进提供了原始材料。

2) 服务层。服务层可以整合数据层的各类数据，通过数据清洗、整合和深入挖掘等技术手段，生成能够刻画设备特征的各类标签。该层致力于实现标签的全周期管理，涵盖了标签的创建、管理、分析和优化四个层面。通过集成先进的数据分析技术，不断对标签体系进行更新和完善，以形成一个全面的设备标签系统。

3) 应用层。基于已有设备画像标签体系，应用层致力于创建满足各种需求的应用场景，包含基于设备画像的供应商评价、基于设备画像的变压器精细化管控策略应用、基于设备画像的变压器运行状态预警、基于标签组合的风险评价等场景应用。

3.3 标签分类

在数据提取的多维视角中，标签的分类涵盖了事实标签、模型标签和预测标签。事实标签可以直接描述设备的基本属性和资料，为我们提供了设备的基础信息。模型标签则进一步通过分析这些数据，生成设备的综合描述，包括其运行和安全状况。而预测标签则基于设备的各种信息，预测潜在风险，为设备的检修提供指导。这三种标签的结合，不仅加深了我们对设备的理解，也增强了资产管理的效率和设备的运行安全。

4 单体设备综合评分方法的制定

为了更好地评测设备画像所展现的设备情况，围绕安全、效能、成本三个维度，选取相关指标因子，构建SEC综合评分模型，通过对资产的安全、效益、周期成本进行综合评价和分析，最终研究形成了一套独特的评分机制。

通过专家法按照经验给各项指标赋予权重值，具体指标及权重情况如下表：

表1 综合评分的指标和权重

序号	一级指标	权重	权重得分	二级指标	权重	权重得分	评分规则
1	安全	60%	60分	单台设备故障停运次数	30%	18	按年度为周期,发生一次则整项归0
				单台设备故障次数	30%	18	该设备未发生故障满分,发生1个故障扣6分,以此类推
				单台设备缺陷数	20%	12	该设备不存在缺陷满分,发生1个缺陷扣4分,以此类推
				设备状态评价	20%	12	正常状态满分,注意状态10分,异常状态7,严重状态4
2	效能	25%	25分	单设备强迫停运次数	40%	10	按年度为周期,发生一次则整项归0
				单设备年均负载率	30%	7.5	0-20%为轻载,扣2分,20-80%为负载率正常区间,不扣分,80-100%为重载,扣3分
				单设备可用系数	30%	7.5	可用系数 = Σ 该设备可用小时 / Σ 该涉笔统计期间小时 可用系数达到或高于公司目标值得7.5分,每低于目标值0.1%扣2分,7.5分扣完为止。
3	成本	15%	15分	初始投资成本	15%	2.25	分值计算方法: 1、先算出同类型设备各阶段的平均成本费用 2、用当前(个体设备的各阶段成本-同类型设备各阶段平均成本)/同类型设备各阶段平均成本,算出比例,按照超出的比例扣除分数 (1)个体设备各阶段成本 < 同类型设备各阶段平均成本,则不扣分 (2)个体设备各阶段成本 > 同类型设备各阶段平均成本,则(个体设备各阶段成本-同类型设备各阶段平均成本)/同类型设备各阶段平均成本,算出比例,按照超出的比例扣除分数。扣完为止 3、算出成本初始投资成本、运维期成本、退役处置成本所有阶段的分数后相加得到成本部分的总得分 注:设备未覆盖的阶段不扣分
				运维期成本	80%	12	
				退役处置成本	5%	0.75	

然后通过归一法将所有指标结果带入统一公式进行计算,得出最终得分(0-100分)。

归一化公式如下:

$$x_{ij} \leftarrow \frac{x_{ij} - \min_i \{x_{ij}\}}{\max_i \{x_{ij}\} - \min_i \{x_{ij}\}}$$

对于负向指标,归一化公式如下:

$$x_{ij} \leftarrow \frac{\max_i \{x_{ij}\} - x_{ij}}{\max_i \{x_{ij}\} - \min_i \{x_{ij}\}}$$

5 结论

设备资产全周期画像总的来说就是由一系列标签构成,本文利用大数据挖掘,以建立的设备画像为基础,按照标签的形式,构建了全方位的设备信息图,从而使商业人员可以迅速获取供应商的信息、变压器的运行情况以及设备的运行潜在风险,更有利于精细化管理和差异化管理。再加上特别的管理评分机制,有助于我们评判设备的具体情况。

通过多级标签体系,综合数据层、服务层和应用层多方面,最终实现设备从数据收集、数据处理到数据应

用的全方位处理,考虑多种因素,充分利用设备画像以及指定的标签特色,最终应用到实际场景中,使我们能够更好的把握设备的运行状态,及时进行管控、优化和检修,在保障设备可靠性和安全性中都会发挥出更加重要的优势。

参考文献

- [1] 邹国惠,魏嘉隆,王超,等.基于知识图谱的变电设备画像技术[J].广东电力,2024,37(01):86-93.
- [2] 张田佳,宦钧,吕晨宇,等.资产全寿命周期管理的信息化体系建设研究[C]//中国设备管理协会.第十二届设备全寿命周期费用技术大会论文集.内蒙古电力集团有限责任公司;2018:5.DOI:10.26914/c.cnkihy.2018.012243.
- [3] 赵永柱,马霖,张可心.基于电力资产全寿命周期的标签画像技术研究[J].电网与清洁能源,2018,34(01):51-58.
- [4] 叶海峰,罗颖婷,许海林,等.基于画像的变压器精细化管理策略研究[J].微型电脑应用,2019,35(02):67-69+95.
- [5] 李鑫,史天运,常宝,等.基于设备画像的机车标签体系构建方法研究[J].交通运输系统工程与信息,2021,21(02):189-195+204.DOI:10.16097/j.cnki.1009-6744.2021.02.027.