

基于专家知识与监测数据联合驱动的高压开关柜状态评估^①

仇翔^{②*} 蒋文泽* 吴麒^{③*} 张宝康* 葛其运**

(* 浙江工业大学信息工程学院 杭州 310014)

(** 科润智能控制股份有限公司 衢州 324100)

摘要 高压开关柜(HVS)作为电力系统的关键设备,对其工作状况进行有效评估可以保障电力系统的安全稳定运行。在工程实践中,由于高压开关柜长期服役于潮湿、高温等恶劣环境下,不可避免的传感器失效或人为因素会导致其设备状态数据存在随机缺失的现象,从而破坏了数据的完整性和可用性,使得对数据质量要求较高的数据驱动方法难以直接用于解决高压开关柜状态评估的问题。为了解决上述问题,研究了一种基于专家知识和监测数据联合驱动的高压开关柜状态评估方法。首先,对高压开关柜系统的内部构成进行了深入分析,并根据区域中设备的功能不同将其分为电缆室、母线室和断路器室三大区域。其次,进一步分析了系统状态、各区域状态及其关键部件状态两两之间的因果关系,从而建立了适用于高压开关柜状态评估的三层贝叶斯网络(BN)拓扑结构。然后,引入专家领域知识设计了适用于高压开关柜系统的3种约束罚函数,并通过求解带有约束的优化问题,改善了不完整数据集下的贝叶斯网络参数估计性能,进而实现了对高压开关柜系统状态的精确评估。最后,在自主设计的10 kV高压开关柜样机上开展了对比验证实验,结果表明,相比于支持向量机(SVM)方法和反向传播(BP)神经网络方法,本文所提方法在状态评估精度上更具优势。

关键词 高压开关柜(HVS); 状态评估; 参数学习; 知识与数据联合驱动; 贝叶斯网络(BN)

高压开关柜是一种用于监控与保护电力线路和用电设备的电力装置^[1],是电力系统中必要的基础部件。由于高压开关柜长期服役于恶劣的环境中,难以避免的绝缘故障或机械故障会导致其工作状态异常,从而影响电力系统的安全可靠运行。根据高压开关柜运行过程中的各项监测数据,对其整机运行状态的健康状况进行评估和等级划分,对于及时发现高压开关柜的潜在安全隐患、降低安全事故的发生概率以及提升电力系统的科学管理水平具有重要的理论和工程意义。

目前,围绕设备运行状态的健康状况评估主要

有基于知识驱动的状态评估方法和基于数据驱动的状态评估方法^[2-3]。其中,前者主要通过机理分析或经验知识建立设备的健康度模型或劣化模型,从而基于模型获得设备的健康状况评估。例如,文献[4]在多旋翼飞机的机理分析基础上,建立了适用于飞机运行状况评估的健康度模型,并通过对多转子传感器异常的实例分析验证了所提方法的有效性。此外,文献[5]考虑了多连杆机械臂受到内部扰动和输出观测扰动的经验知识,设计了基于知识的自适应观测器,从而得到机械臂姿态的实时估计。然而,此类方法需要通过知识建立物理对象的精确

① 国家自然科学基金(62072408)资助项目。

② 男,1980年生,博士,副教授;研究方向:工业数据分析,系统辨识;E-mail: qiuxiang@zjut.edu.cn。

③ 通信作者,E-mail: wq93science@163.com。

(收稿日期:2023-01-04)

数学模型,难以直接适用于物理结构复杂且难以精确建模的高压开关柜系统。

随着传感技术的发展和计算机存储能力的提升,基于数据驱动的方法为解决上述问题提供了新的思路。近年来,依靠大量的运行监测数据和坚实的统计学理论,基于数据驱动的方法被广泛应用于解决高压开关柜的状态评估问题,并取得了阶段性的研究成果。其中,文献[6]通过支持向量机(support vector machine, SVM)拟合局部放电信号与高压开关柜不同状态间的映射关系,实现了对高压开关柜系统绝缘性能的评估。文献[7]利用温升数据表征开关柜的状态,提出一种基于自适应反向传播(back propagation, BP)神经网络的开关柜状态评估方法。然而,高压开关柜的运行模式复杂多变,且反映其健康状况的状态量并不唯一,使得上述只考虑某一类传感信号的状态评估方法难以获得对其全面准确的评估结果。贝叶斯网络(Bayesian network, BN)作为一种新颖的概率图模型,不仅能够实现对多源传感信息的有效融合,而且在不确定知识表达和推理方面具有独特优势,近年来受到了高压开关柜等设备状态评估领域的广泛关注。文献[8]通过构建高压开关柜的BN模型,以图形化的方式描述负荷率、超声波、温湿度等多种传感数据与系统健康状况间的因果关系,得到了较为精确的状态评估结果。此外,文献[9]还考虑了高压开关柜的电压、电流数据以及控制器的位置信号,并通过层次分析法构建BN模型,从而实现了对设备可靠性的定性评估。虽然,上述方法从多源传感数据融合的角度实现了对高压开关柜系统状态的全面分析,并获得了较好的评估性能,但是它们仅考虑了数据完整的情况,存在对设备监测数据的完整性和可用性依赖程度高的问题。

事实上,高压开关柜监测数据的完整性和可用性会受到多方面因素的影响,具体包括:(1)多源传感数据存储过程中由于人为因素造成的数据随机缺失现象;(2)高压开关柜长期服役于潮湿、高温等恶劣环境中,传感器的故障或失效难以避免,使得高压开关柜的多源监测数据存在不完整或部分不可用的问题。通过引入专家经验知识填补缺失数据,提高

监测数据的质量,将有望弥补数据驱动方法对设备监测数据的完整性和可用性依赖程度高的缺陷,进而解决数据缺失下的设备状态评估问题^[10]。然而,现有的文献中尚未涉及将知识驱动的方法和数据驱动的方法以两者联合的形式用于解决监测数据不完整或部分不可用的高压开关柜系统状态评估问题。因此,如何将高压开关柜的专家知识合理融入到数据驱动的状态评估模型中,提出一种适用于缺失数据下高压开关柜系统的综合状态评估方法,并提高其状态评估性能,仍有待进一步研究。

为此,本文提出了一种基于专家知识与监测数据联合驱动的高压开关柜状态评估方法。首先,根据高压开关柜的运行机理将其结构划分为指标层、要素层和目标层,建立其BN拓扑结构。其次,针对监测数据不完整或部分不可用的情况,设计了基于专家知识的BN参数学习算法,构建包含范围约束、不等式约束和加性不等式约束的3种罚函数,并通过求解带有约束的期望最大化(expectation maximization, EM)算法得到参数的最优估计。最后,根据BN拓扑结构和BN参数进行最大后验概率推理,最终得到高压开关柜健康状况的评估结果。本文主要贡献包括以下3个方面:(1)分析了系统状态、各区域状态及其关键部件状态两两之间的因果关系,建立了适用于高压开关柜状态评估的三层BN拓扑结构,从而以图形化的方式描述了高压开关柜各环节间的依赖关系,提高了状态评估模型的可解释性;(2)根据专家知识设计了3种适用于高压开关柜系统的约束罚函数,与文献[11]的方法相比有效提升了高压开关柜监测数据不完整或部分缺失下的模型参数学习效率和估计精度;(3)与SVM、BP神经网络的数据驱动算法相比,本文所提状态评估方法的准确度较高。

1 问题描述与系统建模

1.1 研究对象描述

如图1所示,本文研究建立在与合作单位联合研发设计的10 kV高压开关柜原型样机的基础上。具体来说,样机主要由母线室、断路器室、电缆室以及仪器仪表4块区域构成。图中,母线室布置在高

压开关柜的背面上方,常用于固定三相高压交流母排,并通过支路母线实现与静触头的连接;断路器室主要包括高压负荷开关与高压熔断器,前者专门用于接通和断开负荷电流的电气设备,并通过与高压熔断器串联实现对短路电流的切除;电缆室内安装有电流互感器、接地开关、避雷器(过电压保护器)

以及电缆等附属设备,支持从外部引进电源,以及通过出线电缆连接用电负荷设备;在柜体上方布置了无线测温传感器、弧光传感器、局部放电传感器,从而实现对高压开关柜运行状况的动态监测。为了实现数据上云和远程数据监控的功能,上述仪器仪表均预留了相应的通信转接口。



图 1 10 kV 高压开关柜原型样机及其内部结构示意图

对电力系统中电气部件故障统计数据的研究表明^[12],高压开关柜的异常主要出现在母线室、断路器室和电缆室区域。其中,对于母线室和电缆室,其内部关键部件的绝缘性能是影响其系统状态的主要因素。对于断路器室而言,其状态主要受到内部关键设备的机械性能影响。因此,需要对 3 个区域内

的关键部件进行监测。具体地,针对母线室区域,对其静触头温度、接头温度和母线暂态对地电压进行监测;针对电缆室区域,重点监测其弧光动作次数、电缆温度和电缆室超声波频率;针对断路器室区域,着重监测断路器的合闸时间、分闸时间与动作电流。本文根据如图 2 所示的数据采集系统对 10 kV 高压

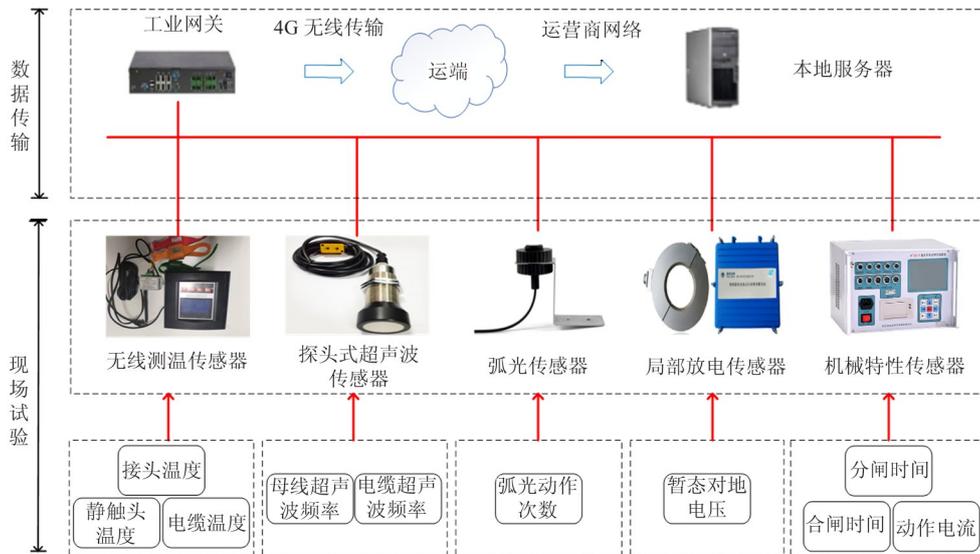


图 2 10 kV 高压开关柜数据采集系统架构图

开关柜原型样机开展数据采集与传输工作,各项监测数据根据《电气装置安装工程电气设备交接试验标准》进行现场试验获得。其中,静触头温度、接头温度和电缆温度数据通过无线测温传感器测量获得;弧光动作次数由弧光传感器采集,能够反映高压开关柜的绝缘性能,其次数越多表明高压开关柜的绝缘劣化程度越高;暂态对地电压数据能够通过局部放电传感器收集;母线超声波频率、电缆超声波频率数据采用探头式超声波传感器分别对母线室和电缆室区域进行测量;合闸时间、分闸时间与动作电流数据通过在真空负荷开关线圈两端接入机械特性传感器进行采集。在数据传输部分,通过工业网关完成多种高压开关柜数据的上传,并采用运营商网络将数据传输至本地服务器,以便后续的状态评估分析。

1.2 问题描述

本文依托 10 kV 高压开关柜数据采集系统收集的各项监测数据,在 BN 模型的基础上对高压开关柜进行状态评估^[12]。BN 是一种概率图形族的网络模型,通过父节点与子节点间的因果关系建模^[13]。在高压开关柜系统中通过静触头温度、接头温度和母线暂态对地电压评估母线室的状态;通过合闸时间、分闸时间与动作电流能够表征断路器室的状态;通过弧光动作次数、电缆温度和电缆室超声波频率能够反映电缆室状态。本文通过上述因果关系,以 3 个区域的状态为父节点,各项监测数据为子节点建立了适用于高压开关柜状态评估的 BN 模型。该 BN 模型定义为 $G = \{V, A, \delta\}$, 其中 V 表示高压开关柜系统变量的集合,且 $x_i \in V (1 \leq i \leq n)$ 表示一个节点; A 表示有向边的集合,网络中的有向边表示节点间的因果关系; δ 表示 BN 模型参数,其量化了节点间的因果关系^[14]。

在基于 BN 的高压开关柜状态评估过程中,往往依赖于完整性和可用性较高的数据。定义长度为 M 、变量种类为 N 的完整高压开关柜数据集 D 为

$$D = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \cdots & x_{1N} \\ x_{21} & x_{22} & x_{23} & \cdots & x_{2N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{M1} & x_{M2} & x_{M3} & \cdots & x_{MN} \end{bmatrix} \quad (1)$$

当高压开关柜数据集完整时, BN 参数通过极大似然估计方法获得,从而评估高压开关柜的状态。

然而,高压开关柜长期服役于潮湿、高温等恶劣环境下,不可避免的传感器失效或人为因素会导致其设备状态数据存在随机缺失的现象。考虑由于传感器失效或人为因素的影响,造成第 $\eta \in (1, 2, \dots, n)$ 个状态数据全部缺失的情况,记为 $[x_{1\eta}, x_{2\eta}, \dots, x_{N\eta}]^T$, 则随机缺失的高压开关柜数据集 \bar{D} 可以表示为

$$\bar{D} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1\eta} & \cdots & x_{1N} \\ x_{21} & x_{22} & \cdots & x_{2\eta} & \cdots & x_{2N} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{M1} & x_{M2} & \cdots & x_{M\eta} & \cdots & x_{MN} \end{bmatrix} \quad (2)$$

如式(2)所示,随机缺失破坏了数据集的完整性和可用性,导致难以获得准确的高压开关柜状态评估模型参数。针对上述问题,如何提高监测数据的质量、获取准确的 BN 模型参数从而解决高压开关柜系统的综合状态评估是本文的关键。

1.3 基于高压开关柜专家知识的约束

为了解决高压开关柜数据的随机缺失,通过引入专家经验知识填补缺失数据,从而提高监测数据的质量。针对高压开关柜系统,其状态数据存在大量的专家先验知识,本文分析设计了 3 种高压开关柜的参数约束。定义 BN 模型参数为 $\delta_{ijk} = P(x_i^k | pa_i^j)$, 其中 P 表示给定节点 $x(i)$ 的父节点取值为 j 时节点 $x(i)$ 取值为 k 的概率, pa_i^j 表示节点 $x(i)$ 的父节点的第 j 个取值, k 表示 BN 节点 x_i 的状态值, i 表示 BN 节点编号, j 表示 BN 父节点的取值。进一步,高压开关柜系统的参数约束可以表示为以下 3 种。

(1) 范围约束

$$\alpha_{ijk} \leq \delta_{ijk} \leq \beta_{ijk} \quad \delta_{ijk} \in C_1 \quad (3)$$

其中, α_{ijk} 表示参数的上界, β_{ijk} 表示参数的下界, C_1 表示所有满足范围约束的参数集合。在高压开关柜系统中,范围约束较为常见,例如不同接头温度下母线室状态处于一个范围内^[15]。

(2) 不等式约束

$$\delta_{ijk} \geq \delta_{i'j'k'} \quad \delta_{ijk}, \delta_{i'j'k'} \in C_2 \quad (4)$$

其中, δ_{ijk} 和 $\delta_{ij'k'}$ 表示一对不同参数, C_2 为所有满足不等式约束的参数集合。例如电缆温度的升高会影响电缆室区域的绝缘性能, 因此电缆温度过高时电缆室区域异常的可能性通常大于电缆温度正常的情况。

(3) 加性不等式约束

$$\sum_{k=1}^{k(c)} \delta_{ijk} \leq \sum_{k=1}^{k(c)} \delta_{ij'k'} \quad \delta_{ijk}, \delta_{ij'k'} \in C_3 \quad (5)$$

其中, $k(c)$ 表示加性不等式约束下节点取值的上限, C_3 表示所有满足加性不等式约束的参数集合。加性不等式描述了高压开关柜 BN 网络中的父节点对子节点的定性影响程度。例如, 父节点断路器的状态取值会直接影响子节点分闸时间的状态^[16], 存在“ $P(\text{分闸时间} = \text{优秀} | \text{断路器状态} = \text{异常}) + P(\text{分闸时间} = \text{良好} | \text{断路器状态} = \text{异常}) \leq P(\text{分闸时间} = \text{优秀} | \text{断路器状态} = \text{优秀}) + P(\text{分闸时间} = \text{良好} | \text{断路器状态} = \text{优秀})$ ”的情况^[17]。

实际运行中, 考虑高压开关柜领域专家受到知识结构、专业水平等因素的影响, 每位专家提出的知识存在不确定性。本文引入置信度 ψ_{ijk} 、 $\psi_{ij'k'}$ 、 $\psi_{ij'k}$ 分别描述以上 3 种约束的不确定性, 其计算公式如下:

$$\psi_{ijk} = \frac{1}{S} \sum_{s=1}^S \psi_{ijk}(s) \quad (6)$$

$$\psi_{ij'k'} = \frac{1}{S} \sum_{s=1}^S \psi_{ij'k'}(s) \quad (7)$$

$$\psi_{ij'k} = \frac{1}{S} \sum_{s=1}^S \psi_{ij'k}(s) \quad (8)$$

其中, S 表示专家的人数, s 表示专家的编号, $\psi_{ijk}(s)$ 、 $\psi_{ij'k'}(s)$ 、 $\psi_{ij'k}(s)$ 分别表示第 s 个专家设置的 3 种置信度。

2 算法设计

围绕高压开关柜监测数据存在随机缺失的问题, 提出了基于专家知识与监测数据联合驱动的高压开关柜状态评估方法, 方法框架如图 3 所示。在 BN 拓扑结构构建阶段, 结合高压开关柜系统的内部结构建立了高压开关柜状态评估的 BN 拓扑结构。进一步, 根据专家知识设计了适用于高压开关柜系统的 3 种约束罚函数, 并通过求解带有约束的 EM 算法优化了参数学习性能, 使得 BN 参数能较快地收敛到最优解, 有效解决数据随机缺失下的 BN 参数训练问题。最后, 在状态评估阶段, 根据训练得到的 BN 参数和 BN 拓扑结构, 采用最大后验概率推理算法获得评估结果。

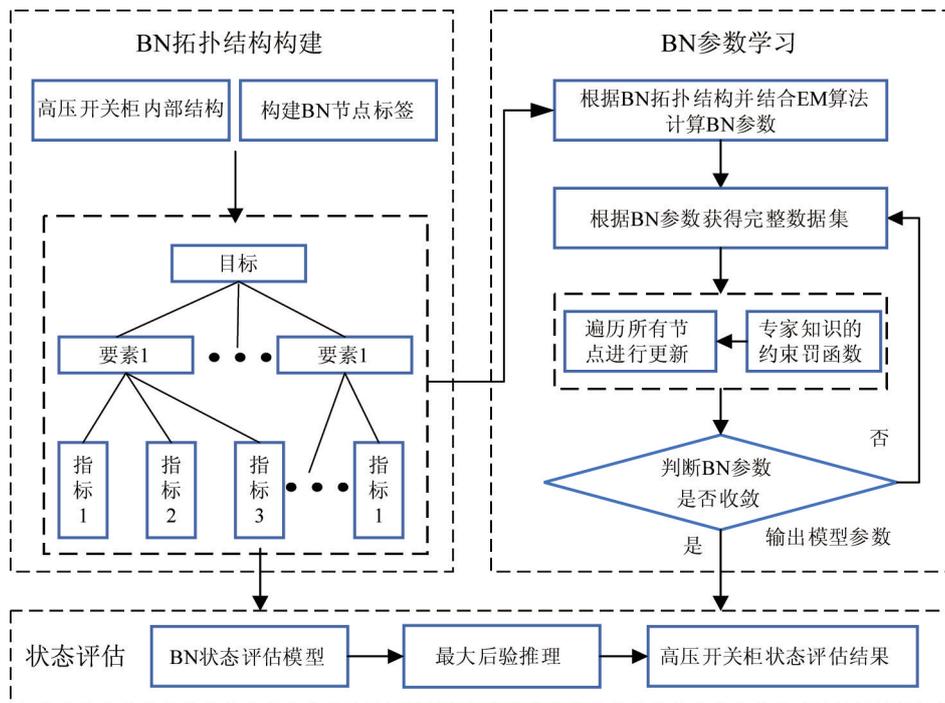


图 3 基于专家知识与监测数据联合驱动的高压开关柜状态评估方法

2.1 高压开关柜状态评估模型构建

结合图 1 高压开关柜的结构特点,本文重点考察了高压开关柜内母线室、断路器室和电缆室 3 个区域的关键传感数据,并根据系统状态、各区域状态及其关键部件状态两两之间的因果关系建立了适用于高压开关柜状态评估的 BN 拓扑结构,如图 4 所示。其中,目标层为高压开关柜系统状态,要素层包括母线室状态、电缆室状态和断路器室状态。指标层选取 3 个区域的关键传感数据。

由于 BN 是一种有监督的学习算法,因此需要对图 3 中的节点构建标签。对于指标层节点根据《10 kV 开关柜技术新要求》,并参考文献[8]中的标准划分健康等级,即赋予优秀、良好、异常 3 种标签,

节点标签如表 1 所示。

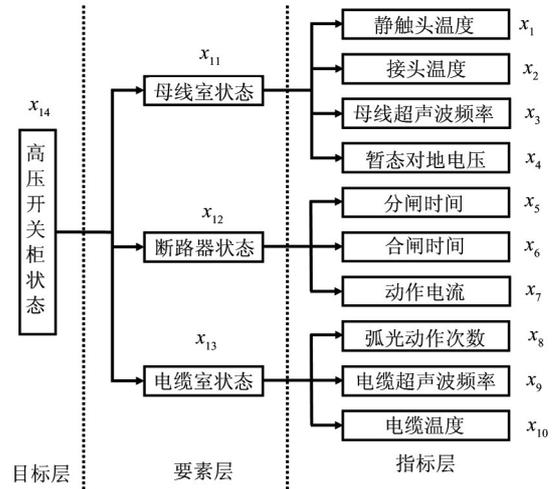


图 4 高压开关柜的 BN 拓扑结构

表 1 节点标签表

节点	节点标签		
	优秀	良好	异常
静触头温度/°C	$x_1 \leq 75$	$75 > x_1 \leq 85$	$x_1 > 85$
母线接头温度/°C	$x_2 \leq 100$	$100 > x_2 \leq 120$	$x_2 > 120$
母线室超声波频率/dB	$x_3 \leq 20$	$20 > x_3 \leq 40$	$x_3 > 40$
暂态对地电压/dB	$x_4 \leq 50$	$50 > x_4 \leq 100$	$x_4 > 100$
分闸时间/ms	$x_5 \leq 50$	$50 > x_5 \leq 60$	$x_5 > 60$
合闸时间/ms	$x_6 \leq 65$	$65 > x_6 \leq 75$	$x_6 > 75$
分合闸电/A	$ x_{10} - 1.8 \leq 0.2$	$0.2 > x_{10} - 1.8 \leq 0.3$	$ x_{10} - 1.8 > 0.3$
弧光动作次数	$x_7 \leq 2$	$2 > x_7 \leq 4$	$x_7 > 5$
电缆超声波频率/dB	$x_8 \leq 25$	$25 > x_8 \leq 45$	$x_8 > 45$
电缆温度/°C	$x_9 \leq 50$	$50 > x_9 \leq 70$	$x_9 > 70$

2.2 结合专家知识的高压开关柜 BN 参数学习

在 2.1 节中,高压开关柜内部结构建立了适用于高压开关柜的 BN 拓扑结构。在此基础上,本小节主要介绍不完整数据下结合专家知识的高压开关柜 BN 参数学习方法。定义高压开关柜 BN 拓扑结构中节点的完整数据为 $D = Y \cup Z$, 其中 $Y = \{Y_l | l = 1, 2, \dots, M\}$ 和 $Z = \{Z_l | l = 1, 2, \dots, M\}$ 分别表示高压开关柜 BN 网络节点的观测数据集和缺失数据集。本文所提 BN 参数学习方法主要包含 3 个步骤:

(1) 在高压开关柜 BN 拓扑结构和当前参数估计下,通过计算高压开关柜的缺失数据集 Z 的条件期望实现数据的补齐。

$$\begin{aligned}
 Q(\delta | \delta^{(t)}) &= E_{\delta^{(t)}}[\log P(D | \delta) | \delta^{(t)}, Y] \\
 &= \sum_z \log P(Y, Z | \delta^{(t)}) P(Z | Y, \delta^{(t)})
 \end{aligned} \quad (9)$$

其中, $P(Z | Y, \delta^{(t)})$ 为观测数据集 Y 和 BN 参数 $\delta^{(t)}$ 下高压开关柜缺失数据 Z 的条件概率分布, $P(D | \delta)$ 为 BN 参数的似然函数。

(2) 通过最大化 $Q(\delta | \delta^{(t)})$, 得到 BN 参数 $\delta^{(t+1)}$ 的更新值^[18]。

$$\delta^{(t+1)} = \arg \max_{\delta} Q(\delta | \delta^{(t)}) \quad (10)$$

(3) 考虑在给定形如式(4)~(6)的高压开关柜专家知识约束下求解 BN 参数 $\hat{\delta}$ 使得高压开关柜完整数据的似然函数最大,则高压开关柜 BN 参数

学习问题可以转化为如下带约束的凸优化问题。

$$\begin{aligned} & \arg \max L_D(\delta) \\ \text{s. t. } & \sum_k \delta_{ijk} = 1 \\ & \alpha_{ijk} \leq \delta_{ijk} \leq \beta_{ijk}, \quad \delta_{ijk} \in C_1 \\ & \delta_{ijk} \geq \delta_{i'j'k} \delta_{ijk}, \quad \delta_{i'j'k} \in C_2 \\ & \sum_{k=1}^{s_i} \delta_{ijk} \leq \sum_{k=1}^{s_i} \delta_{ij'k} \delta_{ijk}, \quad \delta_{ij'k} \in C_3 \end{aligned} \quad (11)$$

针对步骤(3)中带约束凸优化问题,通过构建相应的约束罚函数转化为无约束优化问题。具体地,对式(8)中的高压开关柜专家知识,结合相应的置信度定义如下约束罚函数

$$f_1'(\delta_{ijk}) = [\delta_{ijk} - \alpha_{ijk}]^- \quad \forall \delta_{ijk} \in C_1 \quad (12)$$

$$f_2'(\delta_{ijk}) = [\beta_{ijk} - \delta_{ijk}]^- \quad \forall \delta_{ijk} \in C_1 \quad (13)$$

$$f''(\delta_{ijk}, \delta_{i'j'k}) = [\delta_{ijk} - \delta_{i'j'k}]^- \quad \forall \delta_{ijk}, \delta_{i'j'k} \in C_2 \quad (14)$$

$$f'''(\delta_{ijk}, \delta_{ij'k}) = \left[\sum_{k=1}^{s_i} \delta_{ij'k} - \sum_{k=1}^{s_i} \delta_{ijk} \right]^- \quad \forall \delta_{ijk}, \delta_{ij'k} \in C_3 \quad (15)$$

其中, f_1' 、 f_2' 表示针对范围约束设计的惩罚函数, f'' 表示针对不等式约束设计的惩罚函数, f''' 表示针对加性不等式约束设计的惩罚函数。 $[x]^- = \min(0, x)$ 。

根据式(12)~(15)专家知识的约束罚函数和相应的置信度,原带约束的优化问题可改写为

$$\begin{aligned} J(\delta) = & -\frac{w_1}{2} \sum_{C_1} \psi_{ijk} \left[[f_1'(\delta_{ijk})]^2 + [f_2'(\delta_{ijk})]^2 \right] \\ & -\frac{w_2}{2} \sum_{C_2} \psi_{ijk}^{i'j'k'} (f''(\delta_{ijk}, \delta_{i'j'k}))^2 \\ & -\frac{w_3}{2} \sum_{C_3} \psi_{ijk}^{ij'k} (f'''(\delta_{ijk}, \delta_{ij'k}))^2 + L_D(\delta) \end{aligned} \quad (16)$$

其中 w_i 对应每种高压开关柜专家知识的惩罚权重。

在高压开关柜 BN 参数中,存在固有的公理约束。为消除其对优化问题的影响,采用归一化方法并引入新的参数 ζ_{ijk} 。

$$\delta_{ijk} = \frac{\exp(\zeta_{ijk})}{\sum_{k'} \exp(\zeta_{ijk})} \quad (17)$$

归一化后 ζ_{ijk} 的局部最优解也是 δ_{ijk} 的局部最优解。

结合式(14),并根据链式法则可以得到目标函数的梯度。

$$\nabla_{\zeta_{ijk}} L_D(\delta) = \sum_{l=1}^N P(x_i^k, p\alpha_i^j | D_l, \delta) (1 - \delta_{ijk}) \quad (18)$$

$$\nabla_{\zeta_{ijk}} f_1'(\delta_{ijk}) = \begin{cases} \delta_{ijk}^2 - \delta_{ijk} & \delta_{ijk} \leq \alpha_{ijk} \\ 0 & \text{其他} \end{cases} \quad (19)$$

$$\nabla_{\zeta_{ijk}} f_2'(\delta_{ijk}) = \begin{cases} \delta_{ijk} - \delta_{ijk}^2 & \delta_{ijk} \geq \beta_{ijk} \\ 0 & \text{其他} \end{cases} \quad (20)$$

$$\nabla_{\zeta_{ijk}} f''(\delta_{ijk}, \delta_{i'j'k}) = \begin{cases} \delta_{ijk}^2 - \delta_{ijk} & \delta_{ijk} \leq \delta_{i'j'k}, \quad i \neq i' \text{ or } j \neq j' \\ \delta_{ijk}^2 - \delta_{ijk} - \delta_{ijk} \delta_{i'j'k} & \delta_{ijk} \leq \delta_{i'j'k}, \quad i = i', j = j', k = k' \\ 0 & \text{其他} \end{cases} \quad (21)$$

$$\nabla_{\zeta_{ijk}} f'''(\delta_{ijk}, \delta_{ij'k}) = \begin{cases} \delta_{ijk} - \delta_{ij'k}^2 & \sum_{k=1}^{k(c)} \delta_{ij'k} \leq \sum_{k=1}^{k(c)} \delta_{ijk} \\ 0 & \text{其他} \end{cases} \quad (22)$$

其中,

$$\gamma_1 = \psi_{ijk} [f_1'(\delta_{ijk}) \nabla_{\zeta_{ijk}} f_1'(\delta_{ijk}) + f_2'(\delta_{ijk}) \nabla_{\zeta_{ijk}} f_2'(\delta_{ijk})] \quad (23)$$

$$\gamma_2 = \psi_{ijk}^{i'j'k'} f''(\delta_{ijk}, \delta_{i'j'k}) \nabla_{\zeta_{ijk}} f''(\delta_{ijk}, \delta_{i'j'k}) \quad (24)$$

$$\gamma_3 = \psi_{ijk}^{ij'k} f'''(\delta_{ijk}, \delta_{ij'k}) \nabla_{\zeta_{ijk}} f'''(\delta_{ijk}, \delta_{ij'k}) \quad (25)$$

根据式(20)~(22),目标函数 $J(\delta_{ijk})$ 关于 ζ_{ijk} 的梯度为

$$\begin{aligned} \nabla_{\zeta_{ijk}} J(\delta_{ijk}) = & \sum_{l=1}^N P(x_i^k, p\alpha_i^j | D_l, \delta) (1 - \delta_{ijk}) - w_1 \gamma_1 \\ & - w_2 \sum_{C_2^+} \gamma_2 + w_2 \sum_{C_2^-} \gamma_2 - w_3 \sum_{C_3} \gamma_3 \end{aligned} \quad (26)$$

其中, C_2^+ 为高压开关柜不等式约束中第 1 类的集合, C_2^- 是高压开关柜不等式约束中第 2 类的集合, 且 $C_2^+, C_2^- \in C_2$ 。

根据式(23)目标函数的梯度,采用梯度下降法求解 BN 参数 $\delta_{ijk}^t, \zeta_{ijk}^t$, 其迭代公式如下:

$$\Delta \delta_{ijk}^{t+1} = \Delta \delta_{ijk}^t + P(x_i^k, p\alpha_i^j | D_l, \delta^t) \quad (27)$$

$$\Delta \zeta_{ijk}^{t+1} = \Delta \zeta_{ijk}^t (1 - \zeta_{ijk}^t) + \varphi \quad (28)$$

$$\zeta_{ijk}^{t+1} = \zeta_{ijk}^t + \Delta \zeta_{ijk}^{t+1} \quad (29)$$

其中, $\Delta\delta_{ijk}^0$ 的初值为 0, φ 表示式(23)的后 4 项。

2.3 基于最大后验概率推理的高压开关柜状态评估

在获得 BN 参数后和 BN 拓扑结构后,需要计算高压开关柜状态的概率分布,从而进行准确的状态评估。本文采用最大后验概率推理算法推演网络节点信息,实现高压开关柜的状态评估。将待评估的数据输入到已训练的 BN 模型中,定义其证据节点集合为 $E = \{x_1, x_2, \dots, x_{13}\}$ 。根据贝叶斯公式,在证据节点集合条件下高压开关柜的状态 x_{14} 的概率为

$$P(x_{14} | E, \delta, G) = \frac{P(x_{14}, E, \delta, G)}{P(E, \delta, G)} \quad (30)$$

其中, G 表示高压开关柜 BN 拓扑结构, δ 表示训练得到的 BN 参数。

对于高压开关柜状态集合 $x_{14} = \{1, 2, 3\}$, 根据式(30)可以得到:

$$x_{14} = \underset{x_{14}}{\operatorname{argmax}} P(x_{14} | E, \delta, G) \quad (31)$$

在式(31)中分别计算高压开关柜在优秀、良好、异常 3 种状态的后验概率 $P(x_{14} = \text{优秀} | E, \delta, G)$, $P(x_{14} = \text{良好} | E, \delta, G)$ 和 $P(x_{14} = \text{异常} | E, \delta, G)$, 并根据最大后验概率准则将概率最大的状态作为高压开关柜的评估结果。

综上,基于专家知识与监测数据联合驱动的高压开关柜状态评估方法步骤如下。

步骤 1 通过高压开关柜的运行机理构建三层 BN 拓扑结构。

步骤 2 根据 BN 参数初值 δ^0 和式(9)计算对数似然函数的条件期望,并通过条件期望补全缺失的数据。

步骤 3 根据式(10),最大化似然函数的期望从而获得 BN 参数。

步骤 4 利用梯度下降法求解式(16)中带有高压开关柜专家知识的优化问题。

步骤 5 在遍历 BN 模型所有节点过程中加入式(12)~(15)专家知识的约束罚函数,并通过式(24)~(26)更新 BN 参数。

步骤 6 判断 BN 参数是否收敛。如若收敛输出 BN 参数 δ , 否则重复步骤 2~5。

步骤 7 在获得 BN 参数基础上,根据状态评估模型 G , 进行最大后验概率推理,求解式(31)获得

高压开关柜的状态评估结果。

3 案例分析

3.1 案例数据

本文在研发的高压开关柜样机基础上对所提方法进行了验证。样机模拟了恶劣运行环境下传感器数据缺失的情况,如何在不完整数据集下实现对其精确的状态评估是现有高压开关柜系统所面临共性难题。具体来说,本文通过样机模拟了 3 种随机缺失数据集对所提出方案进行验证。(1)数据集 1:在数据采集过程中模拟了高压开关柜局部放电传感器失效的情况,并收集了样本长度为 800 的随机缺失数据。(2)数据集 2:考虑多种传感器数据缺失的情况,模拟了由于无线温度传感器失效导致的静触头温度、母线接头温度和电缆温度 3 种数据缺失,并收集了样本长度为 800 的随机缺失数据。(3)数据集 3:在数据采集过程中收集不同样本数量的数据集,并模拟高压开关柜局部放电传感器失效的情况。

此外,通过咨询合作单位中具有丰富理论知识、工作背景和相关实践的 5 位专家形成范围约束、不等式约束、加性不等式约束。专家知识的置信度根据式(6)~(8)计算得到,具体如表 2 所示。

表 2 专家知识的约束

约束	置信度	约束	置信度
$0.58 \leq \delta_{1,1,1} \leq 0.81$	0.78	$0.48 \leq \delta_{6,1,1} \leq 0.83$	0.75
$0.31 \leq \delta_{2,2,3} \leq 0.51$	0.75	$0.32 \leq \delta_{6,1,2} \leq 0.43$	0.82
$0.59 \leq \delta_{2,3,3} \leq 0.84$	0.75	$0.58 \leq \delta_{10,1,1} \leq 0.74$	0.87
$0.60 \leq \delta_{5,1,1} \leq 0.78$	0.81	$0.32 \leq \delta_{13,1,2} \leq 0.61$	0.72
$\delta_{1,1,2} \leq \delta_{1,1,1}$	0.87	$\delta_{6,1,2} \leq \delta_{6,1,1}$	0.72
$\delta_{2,1,2} \leq \delta_{2,1,1}$	0.82	$\delta_{7,1,2} \leq \delta_{7,1,1}$	0.81
$\delta_{2,1,2} \leq \delta_{2,1,3}$	0.75	$\delta_{8,1,2} \leq \delta_{8,1,3}$	0.74
$\delta_{1,1,2} \leq \delta_{1,1,1}$	0.78	$\delta_{9,1,2} \leq \delta_{9,1,1}$	0.91
$\sum_{k=1}^2 \delta_{1,2,k} \leq \sum_{k=1}^2 \delta_{1,1,k}$	0.85	$\sum_{k=1}^2 \delta_{3,3,k} \leq \sum_{k=1}^2 \delta_{5,1,k}$	0.74
$\sum_{k=1}^2 \delta_{3,2,k} \leq \sum_{k=1}^2 \delta_{3,1,k}$	0.91	$\sum_{k=1}^2 \delta_{8,3,k} \leq \sum_{k=1}^2 \delta_{8,1,k}$	0.72

3.2 评价指标

为了评估参数学习的效果,本节选择 KL 散度

(Kullback-Leibler divergence) 和负对数似然 (negative log-likelihood, NLL) 作为评价指标。

KL 散度描述了 2 个概率分布间差异的非对称性度量^[19],其绝对值越小,估计的参数越接近真实值。对于模型的真实参数 δ_{ijk} 及其估计值 $\hat{\delta}_{ijk}$, 平均 KL 散度可以表示为

$$\overline{KL}(\delta_{ijk} | \hat{\delta}_{ijk}) = \frac{1}{\sum_{i=1}^n g_i} \sum_{i=1}^n \sum_{j=1}^{g_i} \sum_{k=1}^{s_i} \delta_{ijk} \log \frac{\delta_{ijk}}{\hat{\delta}_{ijk}} \quad (32)$$

负对数似然用以评估学习的 BN 参数与真实数据整体吻合的程度。负对数似然的值越小,学习的贝叶斯网络参数越符合测试数据,其可以表示为

$$NLL(\delta_{ijk} | \hat{\delta}_{ijk}) = - \sum_{i=1}^n \sum_{j=1}^{g_i} \sum_{k=1}^{s_i} \delta_{ijk} \log \hat{\delta}_{ijk} \quad (33)$$

此外,本文选取准确率作为状态评估性能对比指标^[20],具体定义如下:

$$\text{准确率} = \frac{1}{H} \sum_{h=1}^H \frac{b_h}{B_h} \quad (34)$$

其中, H 为高压开关柜状态类别总数, b_h 为第 h 类中评估正确的样本数, B_h 为第 h 类的状态样本总数。

3.3 算法验证及结果分析

首先,为了充分验证所提算法在解决高压开关柜数据缺失情况下 BN 参数学习问题上的收敛性能,采用数据集 1 进行验证。在案例中引入了标准 EM 的参数学习方法作为对照,并随机选取相同的初值,设置最大迭代次数为 20,收敛精度为 0.01。两种参数学习方法的训练结果如图 5 所示。可以发现,2 种参数学习方法最终都能收敛,且具有较低的负似然对数得分,此外,结合高压开关柜的专家约束知识后本文方法在收敛速度上和模型参数估计精度指标上要优于标准 EM 的方法。这是由于专家知识缩小了优化算法的搜索空间,使得算法能够更快收敛到最优值。

其次,本文还考察了不同的高压开关柜传感器数据缺失对参数学习效果的影响。图 6 给出了 2 种算法在数据集 1 和数据集 2 下参数学习性能的对比。结果表明,传感器失效的数量会影响参数学习的效果。当数据集中缺失数据的节点增加时,2 种算法

的精度都有所降低,并且对于缺失节点的参数学习效果要明显差于其他可观测节点。值得注意的是,所提算法相对于基于标准 EM 的参数学习方法有更好的学习性能。

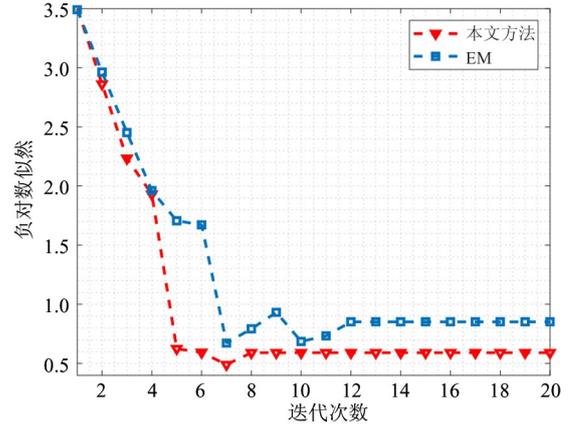


图 5 数据集 1 的学习结果

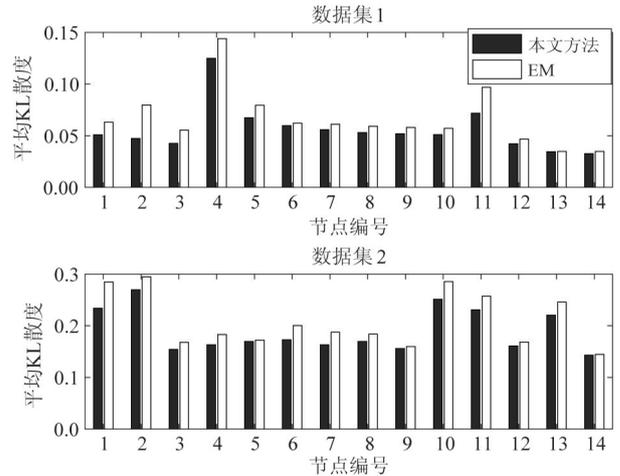


图 6 不同数据缺失下算法的参数学习性能

最后,通过数据集 3 将本文方法所提方法与 SVM、BP 神经网络在状态评估性能方面进行了对比。其中,BP 神经网络的参数配置为 1 个隐藏层,1 个输入层,1 个输出层,激活函数为双曲正切函数,训练误差为 0.01;SVM 的参数配置为惩罚系数为 1.8,核宽系数为 0.2。图 7 为 3 种方法的评估准确率随不同样本长度的变化曲线图,其中评估准确率是实验 50 次的结果。结果表明,随着高压开关柜样本数量的增加,3 种状态评估方法的准确率得到了相应的提升。值得注意的是,与不同的评估方法相比,本文方法一直具有较高的状态评估准确率。

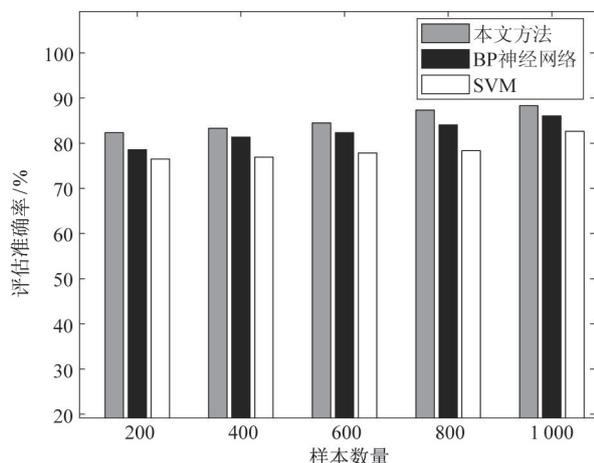


图7 不同样本长度时的状态评估结果

4 结 论

针对高压开关柜数据缺失下的状态评估问题,提出了一种基于专家知识与监测数据联合驱动的方法。通过综合考虑多源因素和系统内部结构的影响,获得了评价高压开关整体状态的三层 BN 拓扑结构,能够显式清晰地描述高压开关柜各环节对其健康状态的影响。尤其针对部分 BN 节点数据缺失下的高压开关柜状态评估问题,提出了基于 EM 算法和专家约束知识相结合的 BN 参数学习方法,优化了贝叶斯网络的模型参数,且不易陷入局部最优。自主设计的 10 kV 高压开关柜样机案例表明,与 EM 算法相比,本文 BN 训练方法具有更好的收敛精度和收敛速度。此外,在案例中将本文方法与基于 SVM、BP 神经网络的评估方法进行对比,结果表明本文所提方法状态评估的准确度较高。

参考文献

- [1] 董盼, 杨鑫, 贾鹏飞, 等. 10 kV 高压开关柜安全性能的提升方法[J]. 电工技术学报, 2022,37(11):2733-2742.
- [2] MENG F, WONG L N Y, ZHOU H. Rock brittleness indices and their applications to different fields of rock engineering: a review[J]. Journal of Rock Mechanics and Geotechnical Engineering, 2021,13(1):221-247.
- [3] WU Q, DONG S, ZHANG W A, et al. Online modeling of the CNC engraving system with dead-zone input nonlinearity[J]. IEEE Transactions on Industrial Electronics, 2021,69(1):774-782.
- [4] ZHAO Z, YAO P, WANG X, et al. Reliable flight performance assessment of multirotor based on interacting multiple model particle filter and health degree[J]. Chinese Journal of Aeronautics, 2019,32(2):444-453.
- [5] DANG S, KONG Z, PENG L, et al. Adaptive state observer for robot manipulators diagnostics and health degree assessment[J]. Applied Sciences, 2020,10(2):514.
- [6] 何金, 田涛, 宋晓博, 等. 基于特高频法的开关柜局部放电识别方法研究[J]. 高压电器, 2020,56(11):90-95,101.
- [7] LIAO Z W, ZHOU K H, ZANG X C. Research on fault diagnosis of switchgear contacts based on BP neural network[C]//2018 International Conference on Power System Technology. Guangzhou, China: IEEE, 2018:3507-3513.
- [8] 亢超群, 李二霞, 盛万兴, 等. 基于多源信息融合的配电真空开关柜动态状态评价研究[J]. 高压电器, 2017,53(3):235-241.
- [9] ZENG Q, CHEN G, HOU H, et al. Reliability assessment on switchgears and its controller based on fault tree analysis[C]//2018 China International Conference on Electricity Distribution. Tianjin, China: IEEE, 2018:972-978.
- [10] LIU Z, LIU Y, ZHANG D, et al. Fault diagnosis for a solar assisted heat pump system under incomplete data and expert knowledge[J]. Energy, 2015,87:41-48.
- [11] KOU L, LIU C, CAI G, et al. Fault diagnosis for open circuit faults in NPC inverter based on knowledge driven and data driven approaches[J]. IET Power Electronics, 2020,13(6):1236-1245.
- [12] 侯中妮, 靳小龙, 陈剑赞, 等. 知识图谱可解释推理研究综述[J]. 软件学报, 2022,33(12):4644-4667.
- [13] 谢修娟, 莫凌飞, 李香菊, 等. 情境感知的移动阅读个性化推荐算法研究[J]. 高技术通讯, 2019,29(7):640-647.
- [14] 张泽鑫, 李俊, 常向青. 基于特征加权的朴素贝叶斯流量分类方法研究[J]. 高技术通讯, 2016,26(2):119-128.
- [15] 董盼, 杨鑫, 贾鹏飞, 等. 10 kV 高压开关柜安全性能的提升方法[J]. 电工技术学报, 2022,37(11):2733-2742.
- [16] 杨帆, 邓一帆, 李东东, 等. 基于最优等级数的多维

- 特征量开关柜健康状态评价方法[J]. 高电压技术, 2020,46(11):3934-3942.
- [17] YANG Y, GAO X, GUO Z, et al. Learning Bayesian networks using the constrained maximum a posteriori probability method[J]. Pattern Recognition, 2019,91:123-13.
- [18] 董青, 郑建飞, 胡昌华, 等. 基于两阶段自适应 Wiener 过程的剩余寿命预测方法[J]. 高技术通讯, 2022,48(2):539-553.
- [19] 闫浩, 王福利, 孙钰洋, 等. 基于贝叶斯网络参数迁移学习的电熔镁炉异常工况识别[J]. 自动化学报, 2021,47(1):197-208.
- [20] 叶远波, 李端超, 谢民, 等. 基于 SSA-SVM 的继电保护装置状态评估方法研究[J]. 电力系统保护与控制, 2022,50(8):171-178.

State evaluation for high-voltage switchgears by combined domain-knowledge-driven and monitored-data-driven methodology

QIU Xiang^{*}, JIANG Wenze^{*}, Wu Qi^{*}, ZHANG Baokang^{*}, GE Qiyun^{**}

(* College of Information Engineering, Zhejiang University of Technology, Hangzhou 310014)

(** Kerun Intelligent Control Co., Ltd., Quzhou 324100)

Abstract

High-voltage switchgear (HVS) is an essential component of the power system, and its state evaluation is of great significance for maintaining stability and safety of the system. In fact, HVS is usually served in such harsh environments with high temperatures and humidity. It is common knowledge that the state of HVS is lost randomly due to unavoidable sensor failure and uncontrollable human factors, which has limitations on the integrity and availability of data. As a result, data driven methods which require high-quality data are hard to be directly applied for state evaluation of HVS. To solve the above problem, an integrated state evaluation approach is proposed by combining domain-knowledge-driven methodology with monitored-data-driven methodology. First, the internal structure of the HVS is in-depth analyzed and it is categorized into three regions pursuant to the terms of the device utility, i. e., the cable-room region, busbar-room region, and circuit breaker-room region. Then, a three-layer Bayesian network (BN) topology for state evaluation of HVS is conducted by the causality analysis of the relationship between each pair of system states, regional states, and basic device states. While the expert domain knowledge is adopted and three kinds of constraint penalty functions matching the BN model are developed. Since the constrained optimization problem is solved, the parameter estimation performance of BN model under incomplete data sets is improved and the accurate state evaluation of the HVS is accomplished. Finally, comparative experiments are carried out on the self-designed 10 kV HVS prototype. The results show that the proposed approach is able to achieve the goal of accurate state evaluation and has superior performance in items of both accuracy and convergence compared with support vector machines (SVM) and back propagation (BP) neural network.

Key words: high-voltage switchgear (HVS), state evaluation, parameter learning, knowledge-driven and data-driven, Bayesian network (BN)