

基于关联规则和多元状态估计的汽轮机故障 预警算法

邹红波, 张馨煜, 李奇隆

(三峡大学 电气与新能源学院, 湖北 宜昌 443000)

摘要:为准确实现汽轮机故障预警,提出一种基于关联规则和多元状态估计的汽轮机故障预警算法。利用关联规则从历史数据中挖掘不同类型的汽轮机故障特征,构建原始决策表;对关联规则挖掘结果展开属性约简,提取最佳故障特征约简集合;采用最佳属性集合组合的约简决策表,对不同类型汽轮机故障信息展开分类。基于汽轮机故障历史数据和实时监测数据组建动态记忆矩阵,使用多元状态估计方法计算汽轮机故障状态,利用非线性欧氏距离计算故障特征向量与观测数据估计偏离距离。利用相似度函数对汽轮机的状态展开预估,设定故障报警阈值,实现汽轮机故障预警。测试结果证明,采用本文算法可准确划分汽轮机故障类型,预估值与实际值基本吻合,可精准实现汽轮机故障预警。

关键词:关联规则;多元状态估计;汽轮机故障预警

中图分类号:TK267 **文献标志码:**A **文章编号:**1671-5497(2024)11-3283-06

DOI:10.13229/j.cnki.jdxbgxb.20230667

Turbine fault warning algorithm based on association rules and multivariate state estimation

ZOU Hong-bo, ZHANG Xin-yu, LI Qi-long

(College of Electrical and New Energy, China Three Gorges University, Yichang 443000, China)

Abstract: In order to accurately achieve steam turbine fault warning, a steam turbine fault warning algorithm based on association rules and multivariate state estimation is proposed. Mining different types of turbine fault features from historical data using association rules to construct original decision tables; Conduct attribute reduction on the results of association rule mining and extract the optimal set of fault feature reduction; Using the optimal attribute set combination reduction decision table to classify different types of turbine fault information. Based on historical data of steam turbine faults and real-time monitoring data, a dynamic memory matrix is constructed, and the fault state of the steam turbine is calculated using multivariate state estimation method. The fault feature vector is calculated using nonlinear Euclidean distance to estimate the deviation distance from the observed data. Use similarity function to estimate the state of the steam turbine, set fault alarm threshold, and achieve steam turbine fault warning. The test results show that the proposed algorithm can accurately classify the types of steam turbine faults, and the

收稿日期:2023-06-28.

基金项目:国家自然科学基金项目(52107108).

作者简介:邹红波(1978-),男,副教授,博士.研究方向:虚拟仪器,信号仿真处理.E-mail:zouhongbo7895@163.com

estimated values are basically consistent with the actual values, which can accurately achieve steam turbine fault warning.

Key words: association rules; multivariate state estimation; steam turbine fault warning

0 引言

电力能源需求日益增加,促使驱动电力机组的汽轮机运行参数越来越高,自动化规模不断扩大,对电厂汽轮机设备运行安全提出了更高的要求,因此,在汽轮机故障萌芽状态或者程度比较轻微时提前预警运行故障,对确保电力系统稳定运行具有十分重要的意义^[1,2]。

相关专家针对故障预警方面的内容展开了大量研究,例如谢天等^[3]全面考虑静态以及动态信息表达形式,将数据特性和机理分析结果有效融合,并且选择满足需求的特征组合,通过长短期记忆网络构建动态预警模型,通过模型实现预警处理;刘朋印等^[4]通过轴系参数和部分未知参数的汽轮机组,分别给出了基于理论分析以及扰动激励辨识的扭矩在线估计方法,通过扭矩评估值评估轴承扭振风险,最终实现在线预警;王宇飞等^[5]通过预警初始时刻的网络攻击观测序列准确辨识各种类型的电网攻击,同时建立对应的预警解集合,利用动态计算的网络安全预测误差和阶段性故障误差设定为预警的判定依据,最终实现故障预警。在上述几种研究算法的基础上,本文提出一种基于关联规则和多元状态估计的汽轮机故障预警算法。通过实验分析证明,采用本文算法可以准确地对汽轮机展开故障分类以及故障预警处理。

1 汽轮机故障预警算法

1.1 基于关联规则的汽轮机故障分类

从电厂监控系统实时数据库中选取汽轮机本体和轴系测点某时间段内的运行数据和电厂汽轮机的历史运行数据。利用关联规则从历史数据中挖掘汽轮机故障属性特性,如压力、振动等,建立汽轮机故障分类模型。

利用关联规则挖掘^[6,7]历史数据中的故障事务数据集。设定 $P = \{p_1, p_2, \dots, p_k\}$ 为汽轮机历史故障数据集。汽轮机故障数据特征关联规则可以表示为 $X \Rightarrow Y$, 其中运行故障数据 $X \in P$, 故障结果数据 $Y \in P$ 。度量关联规则主要包含个比

较重要的指标,分别为:支持度、置信度、相关度。

关联规则 $X \Rightarrow Y$ 的支持度 r 代表汽轮机运行数据集 P 中包含 $X \cup Y$ 的元组所占比例,采用概率论的方式可以将其表示为如下形式:

$$r = \frac{X \cup Y}{P} \quad (1)$$

关联规则 $X \Rightarrow Y$ 的置信度 β 代表 $X \cap Y$ 的数量在 X 元组数量的比例,通过概率论可以将其表示为式(2)的形式:

$$\beta = \frac{X \cap Y}{X} \quad (2)$$

最后一个指标是用来描述 X 和 Y 之间的相关度 τ ,通过概率论可以将其表示为式(3)的形式:

$$\tau = \frac{X \cap Y}{XY} \quad (3)$$

采用关联规则挖掘^[8,9]得到的关联规则存在与故障属性关联性较弱的冗余属性,需要利用权重对其进行属性约简,获得故障属性分类集合,下面给出具体的操作步骤:

(1) 提取待约简属性集特征:

$$Q(P) = \frac{[X \Rightarrow Y]r}{\beta\tau} \quad (4)$$

根据关联规则属性约简布局优化结构,计算多维特征之间的相对权重:

$$\omega = \left(\sum_{c=1}^p \frac{|\tau(X_c, \epsilon_c) - \tau(Y_c, \epsilon_c)|^\beta}{Q(P)} \right) \quad (5)$$

式中: c 为特征维度; ϵ_c 为映射误差; β 为邻域半径变化常量。

基于此,得到的故障属性约简公式为:

$$\begin{aligned} Q(P)_{\omega\kappa\lambda\zeta} &\rightarrow Q(P)_{\omega\lambda(\zeta+1)} \\ Q(P)_{\omega(\zeta+2)} &\rightarrow Q(P)_{\omega(\zeta+l)} \end{aligned} \quad (6)$$

式中: κ, λ, ζ 分别为第一层、第二层、多层的数据总量; l 为决策属性变化量。

根据属性约简结果,结合关联规则进行汽轮机故障分类,步骤如下所示:

(1) 采集汽轮机历史故障信息。

(2) 通过汽轮机历史故障信息建立原始决策表 S , 将各个故障类型对应的汽轮机元件设定为

分类决策属性,组建需要挖掘的故障属性,根据故障属性构建原始决策表。

(3)将获取的全部汽轮机故障数据 P 输入数据库内,构建对应的汽轮机故障数据源,根据构建的数据源组建关联规则 $X \Rightarrow Y$ 挖掘对应的事务表,并且采用 C 语言在 NET 平台提取对应的故障属性集特征。同时利用式(5),选择权重高的属性构建关联频繁项集,并进行约简。

(4)通过约简的关联频繁项集建立更加简单的汽轮机故障决策表 $Q(P)_{\omega(\zeta+t)}$,采用构建的汽轮机故障决策表完成汽轮机故障分类处理,作为后续故障预警的基础依据。

1.2 基于多元状态估计的汽轮机故障预警

基于记忆矩阵建立实时状态数据监测模型,利用多元状态估计^[10]方法对其与上节获得的故障分类特征数据展开最优相似性估计;当汽轮机出现故障时,实时观测数据与上节获得的故障分类特征数据的估计相似性下降,判断汽轮机偏离正常运行状态,进行故障预警。

记忆矩阵是建立多元状态估计模型的重要基础,建立记忆矩阵,设定一个汽轮机共含有 n 个状态参数(压力、振动),则在 t 时刻记录下的监测向量 $Z(t)$ 可以表示为式(7)的形式:

$$Z(t) = [z_1(t), z_2(t), \dots, z_n(t)]^T \quad (7)$$

式中: T 为运行周期。

其中,上节提取的汽轮机故障历史特征数据记忆矩阵 $F_{n \times m}$ 主要是由 m 组观测向量所组成,对应的计算式如式(8)所示:

$$F_{n \times m} = Q(P)_{\omega(\zeta+t)} \begin{bmatrix} f_{1.1}, f_{1.2}, f_{1.3}, \dots, f_{1.m} \\ f_{2.1}, f_{2.2}, f_{2.3}, \dots, f_{2.m} \\ \vdots \\ f_{n.1}, f_{n.2}, f_{n.3}, \dots, f_{n.m} \end{bmatrix}^T \quad (8)$$

其中:矩阵的行是在不同采样时刻下,相同监测向量的变化情况;记忆矩阵的列表代表随机一个采样时刻下,不同监测变量的取值。

(2)分析估计向量和实际向量两者之间的偏差序列,结合记忆矩阵中的不同历史监测向量展开对比分析,进而获取最优估计向量,将其表示为 Z_e ,对应的计算式如式(9)所示:

$$Z_e = \vec{\omega} \cdot F_{n \times m} \quad (9)$$

式中: $\vec{\omega}$ 为权重 ω 的向量,其是通过对历史故障特征矩阵和监测向量加权处理获取的,即估计向量残差平方和,计算式如式(10)所示。

$$\vec{\omega} = \sqrt{F_{n \times m}^2 - Z(t)^2} \quad (10)$$

由于记忆矩阵在点乘之后存在不可逆的情况,通常情况下选择非线性欧氏距离替换点乘积运算。故障特征向量与观测数据估计偏离距离表示为:

$$D = \sqrt{\sum_{i=1}^n ((F_{n \times m})^T - Z_e)^2} \quad (11)$$

取值范围为 $D \in (0, +\infty)$ 。在汽轮机正常运行状态下,新监测向量总是处于故障特征记忆矩阵 $F_{n \times m}$ 所代表的空间范围外,相似度较低。反之,当汽轮机处于故障状态,新的监测向量会逐渐偏向记忆矩阵 $F_{n \times m}$ 所描述的故障工作空间内。

由于评估汽轮机运行状态时,由历史故障特征建立的记忆矩阵对故障预警模型性能的影响比较大。最理想的情况就是选择全部故障运行状态下的汽轮机历史数据建立记忆矩阵,这样能够涵盖汽轮机全部故障运行状态数据。

根据多元状态估计模型可以得到汽轮机实际状态和估计状态两者之间的残差向量,通过对残差向量的分析,可以评估汽轮机的状态符合故障状态的程。通过 D 计算汽轮机状态相似度,将其表示为 Sim ,可以更加直观地评价汽轮机的运行状态,对应的计算式如式(12)所示:

$$Sim = \frac{1}{1 + D} \quad (12)$$

式中: $Sim \in (0, 1)$,取值越接近 0,则说明汽轮机的状态越差;反之,取值越接近 1,则说明汽轮机状态越稳定,越接近正常状态。

为了有效避免偏离距离的统计受到不同因素和环境等方面的影响,提升汽轮机故障预警结果的有效性,需要利用滑动窗口展开数据统计,有效消除不同因素产生的影响,降低误报率,并且可以快速且准确地发现偏差函数的连续变化。利用汽轮机状态相似度构建时间序列,通过式(13)计算滑动窗口内的状态相似度均值:

$$\tilde{U}(Sim) = \frac{1}{N} \sum_{j=1}^N U_j Sim \quad (13)$$

式中: N 为滑动窗口的长度; U_j 为第 j 个监测数据的时间序列。

在上述分析的基础上,需要设定报警阈值 $U_{(y)}$,通常情况下以汽轮机故障状态下相似度的最小值 U_{\min} 作为基准计算,有效减小意外因素产生的影响,对应的计算式如式(14)所示:

$$U_{(y)} = f \cdot (\tilde{U}(Sim) - 3\alpha) \quad (14)$$

式中: α 为标准差; f 为人为设定系数。

假设汽轮机的状态相似度超过设定的阈值 $U_{(y)}$, 则说明汽轮机属于故障状态, 需要马上发出警报, 由此完成汽轮机故障预警。

2 测试分析

为了验证本文基于关联规则和多元状态估计的汽轮机故障预警算法的有效性, 通过 MATLAB 平台展开实验分析。将某电厂作为实验对象, 利用 Datalink 软件与厂内汽轮机监控系统数据库连接, 利用软件 Create PB 在数据库中进行汽轮机历史故障时间段和运行数据的选取。汽轮机型号为 CC25MW。汽轮机实时监测不同测点的分布式环境采用 Cloudera Hadoop 5.0 节点集群框架, 集群通信通过 MPI 库实现, 存储在 HDFS 中。具体环境配置参数如表 1 所示。

汽轮机故障监测预警环境如图 1 所示。

基于以上实验环境采集汽轮机故障实时监测数据。通过本文算法基于历史故障数据, 利用关

表 1 实验环境配置参数

Table 1 Experimental environment configuration parameters

名称	配置参数
硬件环境	Intel i7 八核 CPU
网络环境	DDR 20 GB Infiniband
操作系统	Win 10
节点连接网络	天河-1A
数据集样本大小	35GB
编译器	Ifort V10
仿真平台	MATLAB 2019c
Hadoop	Cloudera Hadop 5.0

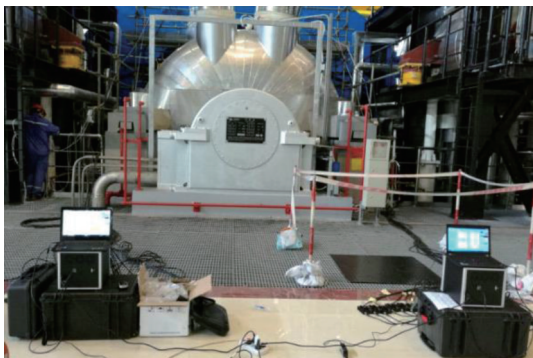


图 1 汽轮机故障实时监测预警环境

Fig. 1 Real time monitoring and warning environment for steam turbine faults

联规则算法进行不同类型的汽轮机故障分类处理, 详细的故障分类结果如表 2 所示。

以表 1 的电厂汽轮机历史故障类型为基准, 从实验环境实时数据库中采集满足相同时间长度的数据, 并对数据进行处理, 输入 MATLAB 平台展开测试。

结合该厂实际情况, 将报警阈值设定为 0.15, 假设汽轮机的状态相似度超过设定的阈值, 则说明汽轮机偏离正常状态, 需要马上发出警报。

选取传统的汽轮机故障分类算法文献[3]方法和文献[4]方法作为对比算法, 3种方法对监测数据的故障分类结果如图 2 所示, 可知, 在不同故障类型下, 各算法对应的汽轮机故障分类精度存在比较明显的差异。和传统的汽轮机故障分类算法相比, 采用本文算法可以准确区分各类型的汽轮机故障, 而传统汽轮机故障分类算法获取的故障分类精度明显比较低。由此可见, 采用本文算法可以获取更高的汽轮机故障分类精度, 为预警奠定了良好基础。

表 2 汽轮机故障分类结果

Table 2 Steam turbine fault classification results

故障类型序号	汽轮机故障类型
01	机组主汽压力故障
02	机组高压缸排汽压力故障
03	轴承盖振动故障
04	机组轴向振动故障
05	机组抗燃油总管压力故障
06	机组主蒸汽压力故障
07	低旁入口蒸汽压力故障
08	机组冷油器出口压力故障
09	机组中压缸压力故障
10	机组低压缸压力故障

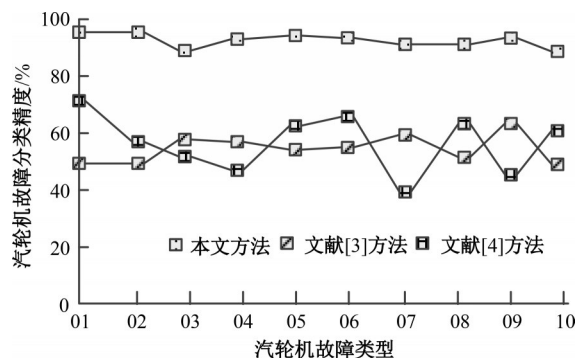
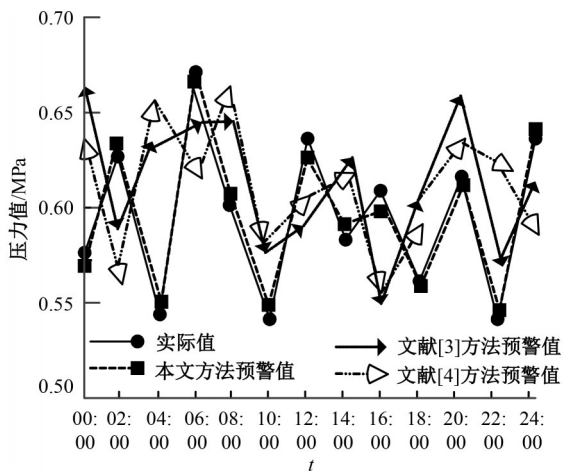


图 2 不同算法的汽轮机故障分类结果比较

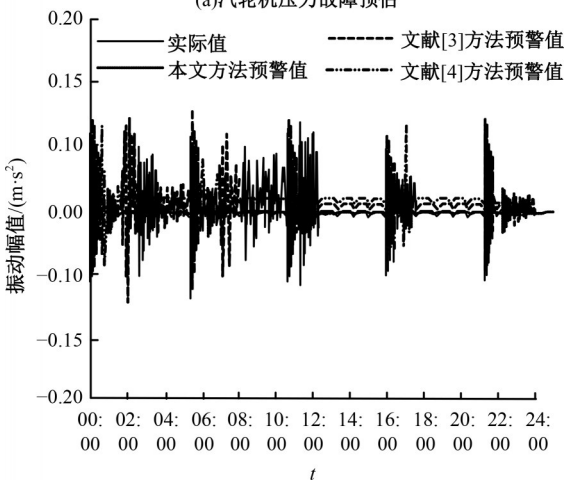
Fig. 2 Comparison of turbine fault classification results using different algorithms

为了进一步验证本文算法的预警性能,以汽轮机抽气压力故障为例进行故障预警展示,3种方法对汽轮机抽气压力的预估值、震动预估值与实际值结果如图 3 所示,可知,与其他方法相比,本文算法获取的压力和振动预估值和实际值基本吻合,误差小于 0.02 MPa 和 0.01 m·s²,说明本文算法可以获得比较精准的汽轮机故障预估结果。

为了进一步验证本文算法的汽轮机故障预警性能,通过均方根误差衡量各个算法预警结果的准确性。其中,均方根误差取值越小,说明汽轮机故障预警结果越准确,详细的实验测试结果如表 3 所示,可知,采用本文算法可以有效减小均方根误差,说明通过本文算法可以获得更加精准的汽轮机故障预警结果。主要是因为本文算法在开展汽轮机故障预警的过程中引入了多元状态估计方法,使其预警性能得到进一步提升。



(a)汽轮机压力故障预估



(b)汽轮机振动故障预估

图 3 汽轮机故障预估仿真结果

Fig. 3 Simulation results of turbine fault warning based on the algorithm

表 3 不同故障类型下各算法的汽轮机故障预警结果比较

Table 3 Comparison of turbine fault warning results for different algorithms under different fault types

汽轮机故障类型	均方根误差/%		
	本文算法	文献[3]算法	文献[4]算法
01	6.384	7.251	7.652
02	5.221	6.558	6.995
03	6.013	6.420	6.784
04	5.447	6.003	6.357
05	6.852	7.123	7.358
06	6.774	7.018	7.280
07	6.890	7.356	7.474
08	6.113	6.458	7.113
09	6.541	6.779	7.056
10	6.101	6.425	6.716

3 结束语

为了有效改善传统汽轮机故障预警算法的不足,提出一种基于关联规则和多元状态估计的汽轮机故障预警算法。通过大量的实验测试证明,采用本文算法可以准确区分各类型的汽轮机故障,同时还可以有效提升汽轮机故障预警结果的准确性,获取更加满意的汽轮机故障预警结果,能有效确保汽轮机的稳定运行。本文方法的局限性在于只显著提升了汽轮机故障预警的准确性,但对汽轮机故障预警效率的提升效果并不明显,未来的研究将在提升汽轮机故障预警准确性的同时对汽轮机故障预警效率进行提高。

参考文献:

[1] 向玲,朱浩伟,丁显,等. 基于CAE与BiLSTM结合的风电机组齿轮箱故障预警方法研究[J]. 动力工程学报, 2022, 42(6): 514-521.
Xiang Ling, Zhu Hao-wei, Ding Xian, et al. Research on fault warning method of wind turbine gearbox based on CAE and BiLSTM[J]. Journal of Chinese Society of Power Engineering, 2022, 42(6): 514-521.

[2] 杨锡运,邓子琦,康宁. 融合集合经验模态分解与宽度学习的齿轮箱故障预警方法[J]. 计算机集成制造系统, 2022, 28(6): 1835-1843.
Yang Xi-yun, Deng Zi-qi, Kang Ning. Early warning method of gearbox fault based on EEMD and broad learning algorithm[J]. Computer Integrated Manufacturing Systems, 2022, 28(6): 1835-1843.

- [3] 谢天, 覃子珍, 杨如意, 等. 基于常模式提取的火电机组通流故障早期预警方法研究[J]. 汽轮机技术, 2023, 65(2): 122-126.
Xie Tian, Qin Zi-zhen, Yang Ru-yi, et al. Study on early warning method for through-flow faults in thermal power units based on constant pattern extraction [J]. Turbine Technology, 2023, 65(2): 122-126.
- [4] 刘朋印, 谢小荣, 马宁宁, 等. 风电次/超同步振荡激发汽轮机轴系扭振风险的在线评估与预警技术[J]. 中国电机工程学报, 2021, 41(增刊1): 52-58.
Liu Peng-yin, Xie Xiao-rong, Ma Ning-ning, et al. Online assessment and early warning of torsional vibration risk for turbine generators stimulated by sub-/super-synchronous oscillations associated with wind power[J]. Proceedings of the CSEE, 2021, 41(Sup. 1): 52-58.
- [5] 王宇飞, 李俊娥, 刘艳丽, 等. 容忍阶段性故障的协同网络攻击引发电网级联故障预警方法[J]. 电力系统自动化, 2021, 45(3): 24-32.
Wang Yu-fei, Li Jun-e, Liu Yan-li, et al. Staged failure tolerance based early warning method for cascading failures in power grid caused by coordinated cyber attacks[J]. Automation of Electric Power Systems, 2021, 45(3): 24-32.
- [6] 辛春花, 郭艳光, 鲁晓波. 大型数据库中利用强化学习改进 treap 的关联规则挖掘算法[J]. 计算机应用研究, 2021, 38(1): 88-92.
Xin Chun-hua, Guo Yan-guang, Lu Xiao-bo. Association rule mining algorithm using improving treap with interpolation algorithm in large database[J]. Application Research of Computers, 2021, 38(1): 88-92.
- [7] 李鑫, 史天运, 常宝, 等. 基于优化的 MsEclat 算法的铁路机车事故故障关联规则挖掘[J]. 中国铁道科学, 2021, 42(4): 155-165.
Li Xin, Shi Tian-yun, Chang Bao, et al. Association rule mining for railway locomotive accident and fault based on optimized MsEclat algorithm[J]. China Railway Science, 2021, 42(4): 155-165.
- [8] 钟倩漪, 钱谦, 伏云发, 等. 粒子群优化算法在关联规则挖掘中的研究综述[J]. 计算机科学与探索, 2021, 15(5): 777-793.
Zhong Qian-yi, Qian Qian, Fu Yun-fa, et al. Survey of particle swarm optimization algorithm for association rule mining[J]. Computer Science and Exploration, 2021, 15(5): 777-793.
- [9] 王培培, 孟芸. 多段支持度数据频繁模式关联规则挖掘仿真[J]. 计算机仿真, 2021, 38(5): 282-286.
Wang Pei-pei, Meng Yun. Simulation of mining frequent pattern association rules of multi-segment support data[J]. Computer Simulation, 2021, 38(5): 282-286.
- [10] 胡杰, 唐静, 谢仕义. 基于实时动态基线的运行设备多元状态估计方法[J]. 热力发电, 2021, 50(2): 125-131.
Hu Jie, Tang Jing, Xie Shi-yi. Multivariate state estimation technique for equipment running condition using real-time dynamic baseline[J]. Thermal Power Generation, 2021, 50(2): 125-131.