

基于数据挖掘的电力系统故障自动诊断方法研究

励力帆,虞伟,桑清城

(国网舟山供电公司,浙江 舟山 316000)

摘要:为了解决当前电力系统故障诊断方法存在的问题,提高故障诊断正确率,提出基于数据挖掘的电力系统故障自动诊断方法。采集电力系统故障数据,从中提取故障诊断特征,对故障诊断特征进行约简处理,得到最优的故障诊断特征。建立电力系统故障诊断的关联规则,根据关联规则实现故障状态诊断,与其他方法进行故障诊断仿真对比实验,结果表明:该方法改善了电力系统故障诊断,而且抗干扰性极强。

关键词:数据挖掘;电力系统;故障自动诊断;粗糙集;免疫算法;关联规则

中图分类号:TM76 **文献标志码:**B **文章编号:**1671-5276(2022)02-0228-04

Research on Automatic Fault Diagnosis of Power System Based on Data Mining

LI Lifan, YU Wei, SANG Qingcheng

(State Grid Zhoushan Power Supply Company, Zhoushan 316000, China)

Abstract: In order to solve some problems existing in current fault diagnosis methods of power system and improve the accuracy of fault diagnosis of power system, an automatic fault diagnosis method of power system based on data mining is proposed. The fault data of power system is collected, and the fault diagnosis features of power system are extracted from the data. The fault diagnosis features of power system are reduced to get the optimal fault diagnosis features of power system. The association rules of power system fault diagnosis are established, and the fault state diagnosis is realized according to the association rules, and the fault diagnosis of power system is carried out with other methods. The simulation results show that the proposed method improves the power system fault diagnosis and has strong anti-interference.

Keywords: data mining; power system; fault automatic diagnosis; rough set; immune algorithm; association rules

0 引言

步入 21 世纪后,电力系统设备状态监测变成电力管理的核心问题。伴随电网规模逐渐变大,电力自检性能被看作必备功能^[1]。自动故障诊断性能具有核心作用,如果电网出现故障,电力调度中心将获取海量故障信息。怎样在大量故障数据里高精度获取故障信息,并实现电力故障自动诊断,对电力安全有着不可忽视的作用^[2]。近年来,中国电力企业在设备运行状态监测问题中进行大量研究与创新,全面、高精度获取电力状态信息对保护电力运行安全存在直接影响^[3]。

数据挖掘技术是近几年出现的智能技术,其属于人工智能和数据库技术相融的结晶。通过数据挖掘相关方法可在海量数据中获取潜在的有价值信息。数据挖掘方法目前层出不穷,例如粗糙集方法、模糊集方法、神经网络方法等,但单一数据挖掘方法在诊断电力故障时,容错性能未能符合诊断需求,且诊断时易进入局部最优,诊断速度较慢^[4]。

为了获得更好的电力系统故障自动诊断结果,本文将数据挖掘中的粗糙集方法(rough sets, RS)与免疫算法(immune algorithm, IA)相结合,并结合关联规则数据挖掘

方法,提出基于数据挖掘的故障自动诊断方法,先实现故障点定位,再针对故障点实现故障状态诊断。

1 数据挖掘的电力系统故障自动诊断方法

1.1 基于数据挖掘的故障定位方法

基于数据挖掘的故障定位方法流程如下:

- 1) 使用电力信息系统有关知识,建立故障挖掘的数据库^[5];
- 2) 获取故障特征,设置和其对应的条件属性与决策属性;
- 3) 按照第 2) 步所设置条件属性和决策属性,以 RS 决策表形式描述故障集合;
- 4) 将决策表里求约简问题变换为求区分矩阵最小约简数问题,使用 IA 实现最优属性约简;
- 5) 在最优属性约简集合里获取所需故障点定位规则。

最优属性约简步骤如下所述。

Step1: 运算电力系统故障特征决策属性 E 对条件属性 D 的依赖水平 S_D , 并设置 $D_{ore}(D) = \Phi$, 分别删除某个单

重属性 $b \in D$ 。如果 $S_{D-b} \neq S_D$ (S_{D-b} 代表去除单重属性 b 后,故障特征决策属性 E 对条件属性 D 的依赖水平),那么 $D_{ore}(D) = D_{ore}(D) \cup b$, $D_{ore}(D)$ 代表核。如果 $S_{D_{ore}} = S_D$,那么最优属性约简值是 D_{ore} ,反之实施第 2)步。

Step2:建立初始抗体集合并实施编码。通过二进制编码方法,将条件属性 D 数目设成抗体长度,抗体条件属性是否有效需要根据基因判断^[6]。约简过程中留下某条件属性时,则抗体基因值必须是 1;若基因值是 0,则对应的条件属性需要去除。初始化过程中,将核里的条件属性值设成 1,抗体规模设成 M ,以此获取初始抗体集合。

Step3:运算亲和度。抗原和抗体间亲和度可以体现可行解是否符合问题要求。亲和度较大,则解符合程度越大^[7-8]。将适应度函数倒数设成亲和度函数,则亲和度运算方法是:

$$g_{axv} = B \times \frac{M - u_v}{M} + (1 - B) \times R \quad (1)$$

式中:条件属性数目设成 M ;抗体 u 中“1”的数目设成 u_v ;依赖水平与调整系数依次设成 R 、 B 。

Step4:运算抗体浓度。先运算两个抗体间亲和度:

$$g_{ayvw} = \frac{1}{1 + e_{\text{fferw}}} \quad (2)$$

式中抗体间结合水平表示同一方位基因编码值存在差异的数目,此值设成 e_{fferw} 。

那么抗体 u 在种群中浓度是:

$$S = \frac{1}{M} \sum_{\omega=1}^M g_{acvw} \quad (3)$$

$$g_{ayvw} = \begin{cases} 1 & g_{ayvw} > H_{\text{acl}} \\ 0 & \text{其他} \end{cases} \quad (4)$$

式中: H_{acl} 为阈值; ω 为权值。

Step5:抗体控制与优化。为了提升抗体的多样化,需要优化亲和度较高的抗体浓度。但是,抗体浓度超标,亲和作用可能被抑制,需要掌握抗体控制与优化的程度^[9]。

Step6:更新记忆库。把每个抗体群里亲和度高、浓度小的 z 个抗体保存在记忆库里,并持续更新^[10]。

Step7:下一代父代抗体群(后续故障点)主要通过选择、交叉与变异形式获取。

Step8:符合停止条件便停止,输出故障点定位结果;反之回到 Step3 步再次循环操作。

综上所述,根据最优属性约简值能够获取决策规则,可把故障和决策规则相融,快速定位电力故障点^[11-12]。

1.2 基于数据挖掘的故障自动诊断方法

将 1.1 小节定位的故障点故障信息设成 $A = (A_1, A_2, \dots, A_n)$,故障区域是 $W = (W_1, W_2, \dots, W_m)$,故障诊断关联规则为 $I \Rightarrow J$,当中 I 、 J 依次代表 A 、 W 的故障模式。

基于关联规则数据挖掘电网故障诊断方法的诊断流程图见图 1。首先将电力故障点信息作为基于关联规则数据挖掘的故障自动诊断方法的输入,按照故障点信息构建故障状态决策表,在此表中设置故障状态关联频繁项,分析故障状态关联频繁项的关联规则,此规则需要按照电力故障点实际状况实施设置,若规则符合实际需求,便通

过规则内容实现故障状态诊断。

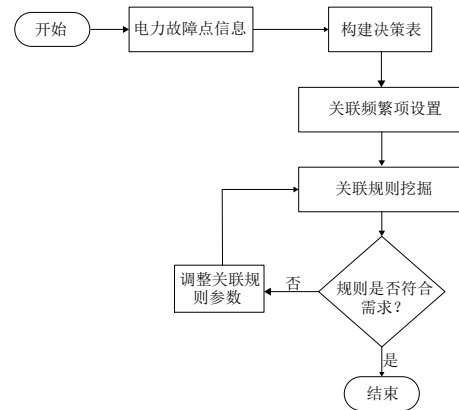


图 1 故障自动诊断流程

电力系统里电力数据存在巨量性,为了详细描述图 1 中的诊断过程,图 2 电力系统为例进行详述。如图 2 所示,此电力系统存在 3 个线路,3 个线路均存在过流保护装置 DP1、DP2、DP3。在线路 1 中设置距离保护装置 SS1,用来实现线路 2 与线路 3 的后备保护。DC1、DC2、DC3 为断路器。将故障诊断时的条件属性值依次设成 1 和 0,若故障诊断时的条件属性值是 1,则断路器从闭合变成断开或保护状态,若属性值是 0,则断路器没有出现变位。

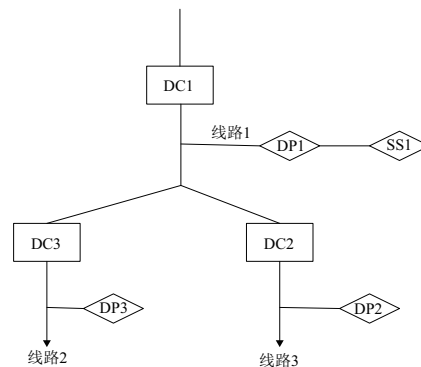


图 2 电力系统结构

将故障信息设成断路器动作信息,将决策属性设成故障点中的故障状态。将故障点设定 6 种故障诊断的关联规则。

规则 1:若故障点出现在线路 1,运行过流保护装置 DP1,跳开断路器 DC1,此线路短路;

规则 2:若故障点出现在线路 2,运行过流保护装置 DP2,跳开断路器 DC2,此线路短路;

规则 3:若故障点出现在线路 3,运行过流保护装置 DP3,跳开断路器 DC3,此线路短路;

规则 4:若故障点出现在线路 2,运行过流保护装置 CP2,而 DP3 断路器没有开启,后备距离保护 SS1 运行,跳开断路器 DC1,线路 2 短路;

规则 5:若故障点出现在线路 3,运行过流保护装置 CP3,而断路器 DP3 未开启,后备距离保护 SS1 运行,跳开断路器 DC1,线路 3 短路;

规则 6:若故障点先出现在线路 2,再出现线路 3,过流保

护失效,后备保护 SS1 开启,跳开断路器 DC1,则双线短路。

2 诊断性能分析

假定图 2 中存在 8 个故障点,详情见图 3。

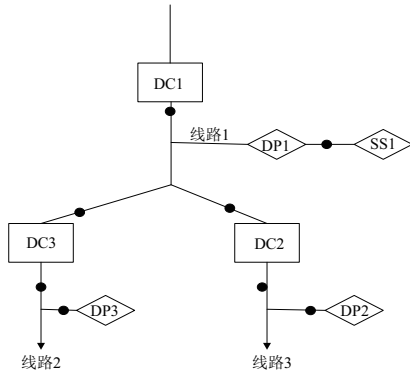


图 3 需诊断的电力故障点示意图

采用本文方法对图 3 中 8 个故障点的故障状况进行诊断。诊断时,使用本文方法获取的最优条件属性约简集合见表 1。

表 1 最优条件属性约简集合

输电网	最优条件属性约简	输电网	最优条件属性约简
故障点 1	D ₂ D ₃ D ₄	故障点 5	D ₂ D ₃ D ₄ D ₅ D ₆ D ₇ D ₈
故障点 2	D ₂ D ₃ D ₄ D ₅	故障点 6	D ₂ D ₃ D ₄ D ₅ D ₆ D ₇
故障点 3	D ₂ D ₃ D ₄ D ₅ D ₆	故障点 7	D ₂ D ₃ D ₄ D ₅ D ₆
故障点 4	D ₂ D ₃ D ₄ D ₅ D ₆ D ₇	故障点 8	D ₂ D ₃ D ₄ D ₅

分析本文方法对 8 个故障点故障状态的误诊率与漏诊率,结果见表 2。分析表 2 可知,本文方法对该大型配电网输电线路中 8 个故障点的误诊率与漏诊率极小,符合诊断需求。

表 2 本文方法误诊率与漏诊率

输电网	漏诊率	误诊率
故障点 1	0.02	0.09
故障点 2	0.01	0.08
故障点 3	0.02	0.09
故障点 4	0.01	0.08
故障点 5	0.02	0.09
故障点 6	0.01	0.08
故障点 7	0.02	0.09
故障点 8	0.01	0.08

为了测试本文方法对电力故障的诊断性能,设定某大型配电网电力系统故障类型依次是单线短路、双线短路、保护或断路器失效、单相接地、开路。测试本文方法、基于故障录波信息的电网故障诊断方法、基于灰色关联度的电网故障诊断方法对 5 种故障的定位精度,结果见图 4。分析图 4 可知,本文方法对 5 种电力故障的定位精度高于对

比方法,且故障定位精度值高达 99.56%。

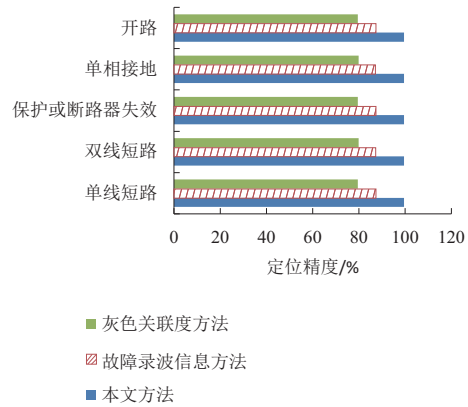


图 4 3种方法故障定位精度对比结果

测试 3 种方法对 5 种电力故障的诊断耗时,结果见表 3。分析表 3 中数据可知,本文方法诊断 5 种电力故障的平均耗时少于两种对比方法且诊断耗时最短,诊断效率最高。

表 3 3种方法故障诊断耗时对比结果 单位:s

类别	本文方法	故障录波信息方法	灰色关联度方法
单线短路	1.11	4.34	7.67
双线短路	1.22	4.33	7.77
保护或断路器失效	1.25	4.35	7.67
单相接地	1.27	4.44	7.77
开路	1.09	4.56	7.77
时间消耗平均值	1.18	4.40	7.73

在 MATLAB 仿真环境中,测试本文方法在实施电力系统故障自动诊断时其生命周期变动情况。设定该电力系统存在 200 个故障点,把各个故障点看成一个节点,分别采用本文方法、基于故障录波信息的电网故障诊断方法、基于灰色关联度的电网故障诊断方法对 200 个故障点进行故障诊断时对电力系统生命周期的影响,结果见图 5。分析图 5 可知,本文方法实施电力系统故障诊断后,电力系统中死亡节点数最大值为 9 个,而其他的对比方法应用后,电力系统中死亡节点数最大值均高于 40 个,和本文方法相比,差异显著。由此可知,本文方法的使用对电力系统生命周期不存在较大损耗,不会干扰电力系统的使用寿命。

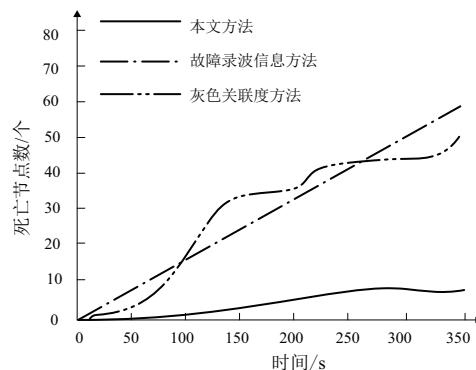


图 5 生命周期变动状况

为了测试本文方法在诊断电力故障时的抗干扰性,在电力系统数据中引入噪声数据,测试3种方法在噪声数据干扰下,诊断电力短路故障时的鲁棒性,结果见表4。表4数据显示,在电力系统数据中引入噪声数据后,伴随噪声数据数目的增多,本文方法的鲁棒性并未受到不良影响,鲁棒性始终 >0.9800 。另外两种方法的鲁棒性虽然也未出现较大变动,但低于本文方法。

表4 3种方法鲁棒性测试结果

噪声数据量/个	本文方法	故障录波信息方法	灰色关联度方法
20	0.989 9	0.910 2	0.892 3
40	0.987 8	0.912 1	0.889 9
60	0.987 7	0.912 3	0.893 3
80	0.989 9	0.912 1	0.893 4
100	0.987 8	0.912 3	0.894 3
120	0.989 9	0.912 1	0.894 5
140	0.987 8	0.912 3	0.893 4
160	0.989 9	0.912 1	0.894 3
180	0.987 8	0.912 3	0.894 3
200	0.988 8	0.910 1	0.897 6

3 结语

本文提出一种基于数据挖掘的电力系统故障自动诊断方法,并在实验中验证其有效性,经验证结论如下:

- 1)本文方法对某大型配电网输电线路中8个故障点的误诊率与漏诊率极小;
- 2)本文方法对5种电力故障的定位精度高于对比方法,故障定位精度值高达99.56%;
- 3)本文方法诊断5种电力故障的平均耗时仅有1.18 s;
- 4)本文方法的使用对电力系统生命周期不存在较大损耗,不会干扰其使用寿命;

5)在电力系统数据中引入噪声数据后,本文方法的鲁棒性并未受到不良影响。

综上所述,本文方法对电力系统故障诊断问题有使用价值。

参考文献:

- [1] 和敬涵,罗国敏,程梦晓,等.新一代人工智能在电力系统故障分析及定位中的研究综述[J].中国电机工程学报,2020,40(17):5506-5516.
- [2] 杨挺,赵黎媛,王成山.人工智能在电力系统及综合能源系统中的应用综述[J].电力系统自动化,2019,43(1):2-14.
- [3] 尚慧玉,王崇宇,陈明辉,等.灾害天气条件下电力系统故障诊断特征匹配方法[J].电力建设,2018,39(10):37-43.
- [4] 张岩,文福拴,熊文,等.基于广域测量信息的电力系统故障诊断解析模型[J].电力建设,2018,39(5):10-20.
- [5] 王磊,陈青,高洪雨,等.基于大数据挖掘技术的智能变电站故障追踪架构[J].电力系统自动化,2018,42(3):84-91.
- [6] 刘科研,吴心忠,石琛,等.基于数据挖掘的配电网故障风险预警[J].电力自动化设备,2018,38(5):148-153.
- [7] 唐远翔,倪秋萍.基于数据挖掘的船舶汽轮机故障诊断系统研究[J].舰船科学技术,2017,39(12):159-161.
- [8] 林济铿,任怡睿,闪鑫,等.基于Logistic回归深度神经网络的电力系统故障概率诊断[J].天津大学学报(自然科学与工程技术版),2021,54(2):186-195.
- [9] 屈子程,高亮,康保林,等.基于多源数据的电力系统故障全信息诊断模型[J].电力系统保护与控制,2019,47(22):59-66.
- [10] 李俊武,李国宁.RS-BN算法在高铁车载设备故障诊断中的应用研究[J].铁道标准设计,2018,62(10):176-182,195.
- [11] 李沛洁,杨博,李宏光.基于关联规则的条件状态模糊Petri网及其在故障诊断中的应用[J].化工学报,2018,69(8):3517-3527.
- [12] 夏昌浩,胡爽,李伶俐,等.基于贝叶斯-粗糙集理论的诊断模型及诊断方法的应用[J].三峡大学学报(自然科学版),2020,42(5):82-87.

收稿日期:2020-10-27

(上接第206页)

- [5] 栗辉.基于iGPS和激光雷达的飞机水平测量设计与实现[D].成都:电子科技大学,2019.
- [6] ZENG Q, HUANG X, LI S G, et al. High-efficiency posture prealignment method for large component assembly via iGPS and laser ranging [J]. IEEE Transactions on Instrumentation and Measurement, 2020, 69(8):5497-5510.
- [7] ZHANG H, PAN X, BAI J P, et al. Research on the application of indoor GPS in aircraft assembly [C]//International Conference on Optical Instruments and Technology 2019: Optoelectronic Measurement Technology and Systems. Beijing: [s.n.], 2020.
- [8] SCHMITT R, NISCH S, SCHÖNBERG A, et al. Performance evaluation of iGPS for industrial applications [C]//2010 International Conference on Indoor Positioning and Indoor

Navigation. Zurich, Switzerland: IEEE, 2010:1-8.

- [9] ZHOU N, AN Z Y, LI L J, et al. iGPS measurement network multi-station arrangement design [J]. Applied Mechanics and Materials, 2013, 443:223-227.
- [10] 孙涪龙,赵罡,王伟,等.iGPS测量不确定度空间分布分析方法[J].北京航空航天大学学报,2015,41(1):174-180.
- [11] 郑迎亚,郝继贵,薛彬,等.室内空间测量定位系统网络布局优化[J].光电工程,2015,42(5):20-26.
- [12] 熊芝,岳翀,薛彬.基于IAGA的空间测量定位系统测站优化部署[J].应用光学,2016,37(4):561-566.
- [13] 岳翀,熊芝,薛彬.基于模拟退火-粒子群算法的wMPS布局优化[J].光电工程,2016,43(7):67-73.

收稿日期:2020-12-24