

基于 GOA-kmeans 的电缆早期故障诊断技术研究

唐辰旭¹, 吉畅²

(1. 国网四川省电力公司德阳供电公司, 四川 德阳 618000; 2. 六盘水师范学院, 贵州 六盘水 553000)

摘要: 电缆早期故障是一种间歇性故障, 持续时间短且不触发继电保护动作, 是永久性故障的前期过程。为识别早期故障的种类和初步估算距离永久性故障的时间, 利用蝗虫优化算法及 k-means 的早期故障识别方法进行电缆早期故障诊断。首先利用蝗虫算法对 k-means 聚类进行改进, 解决 k-means 聚类初始中心点的选择问题, 再基于优化后的 k-means 算法对早期故障波形持续时间、波形特征等完成分类预测, 从而完成早期故障识别。采用实际数据进行验证, 结果表明该方法能够有效识别配电网架空线早期故障, 具有较高的工程应用价值。

关键词: 电缆早期故障; 特征; 蝗虫算法; 实际验证

中图分类号: TM247; U226.8+1; A715

文献标识码: B

文章编号: 1001-2184(2022)05-0092-05

Research on Cable Early Fault Diagnosis Technology Based on GOA-Kmeans

TANG Chenxu¹, JI Chang²

(1. State Grid Sichuan Province Power Co. Ltd., Deyang Power Supply Branch, Deyang Sichuan 618000;
2. Liupanshui Normal University, Liupanshui Guizhou 553000)

Abstract: The early fault of cable is a kind of intermittent fault, which has short duration and does not trigger relay protection action, and is the early process of permanent fault. In order to identify the kinds of early faults and preliminarily estimate the time from early faults to permanent faults, the Grasshopper Optimization Algorithm (GOA) and early fault identification method based on K-means is used to diagnose early faults of cables. Firstly, Grasshopper Algorithm is used to improve the K-means clustering to solve the problem of selecting the initial center of K-means clustering, and then the early fault feature is classified based on the optimized K-means Algorithm, according to the waveform duration and waveform characteristics of the early fault, then the early fault is identified. The results show that the method can effectively identify the early faults of overhead lines in distribution network and has high engineering application value.

Key words: Early fault of cable; Characteristics; Grasshopper Optimization Algorithm; Verification with real data

0 引言

设备的运行过程中并不是所有故障都是瞬时发生的, 我们将正常运行状态至故障状态之间的发展阶段称为早期故障^[1]。早期故障通常来说是由于绝缘损坏、设备老化等产生过程, 一般不会立刻触发继电保护动作, 却会对系统构成安全隐患。因此, 对电力系统早期故障的检测可提高系统可靠性、预防严重故障和降低事故风险具有重要意义。现阶段早期故障诊断技术在电机轴承寿命预测中已经得到较好的应用^[2], 但对配网电缆的早期故障预测尚不成熟。

早期故障通常具备自恢复性、隐蔽性和较强的随机性, 影响因素多, 表现形式为发生的概率

低、信号较弱、持续时间短, 且影响存在故障的电网状态和拓扑在不同时刻性质也不同。如网架结构、线路参数、负载、传感器参数以及噪声等不尽相同, 造成同一类型的故障波形存在较强的随机性。为解决数据的随机性, 目前常采用基于数据驱动的神经网络模型进行数据处理。其中基于傅里叶变换、小波变换、希尔伯特变换的方法是暂态特征提取的有效方法^[3], 利用灰度处理、随机森林、卷积神经网络等方法进行分类是故障特征分类的有效方法^[4-6], 但电缆早期故障的发生频率较低, 数据量较小, 上述方法在实际运用中成功率不高。

本文提出一种利用蝗虫算法(Grasshopper Optimization Algorithm, GOA)改进 k-means 算法进行故障特征分类, 最终达到电缆早期故障诊

收稿日期: 2022-06-17

断的方法。该方法利用蝗虫算法较高的搜索效率、较快的收敛速度及算法本身特殊的自适应机制优化 k-means 算法,解决分类初始点设置的问题。

1 早期故障分类与特征

1.1 早期故障分类

按照故障时暂态信号的来源和持续时间,我们将早期故障分为 4 类。

第一类:暂态干扰(transient disturbance, TD)。该类特征为无序性,故障前后及故障期间都存在,可以通过前后相减去除部分白噪声信号。

第二类:单周波早期故障(subcycle incipient faults, SIF)。单周波早期故障中,电缆电压、电流在对应周波内发生较大的波动,且波动持续半个至一个周波,该类故障主要发生故障最早期的,同一位置连续两次单周波故障之间可能交叉着其他位置的不同故障。

第三类:多周波故障(multicycle incipient faults, MIF)。多周波故障主要发生在永久性故障的临近阶段,特征是电压、电流暂态波形持续时间为 2 个及以上周期,但并未发生永久性的短路、断路等故障,即多周波故障后的电压电流波形与故障前一致,这是区分多周波故障和永久性故障的最大特征。多周波故障发生后往往距离发生永久性故障的时间较短。

第四类:永久性故障(permanent faults, PF)。永久性故障即常见的故障,主要是短路、断线等。永久性故障发生前后电压、电流特点发生了极大改变,因此,可以根据这个特征进行永久性故障的挑选。

1.2 早期故障的特征

综上所述,可以得到早期故障的基本时序特征,以这个基本特征作为参考进行分类,可准确判断属于第几类故障,且根据同类故障的发展趋势,可以初步预估发生的早期故障距离该类永久性故障的时间。但由于不同种类故障早期故障的原因和造成的故障波形不同,对应的故障发展周期就不一致,因此,还需同时对故障波形进行分类,本文采用余弦相似度进行判定。

2 基于蝗虫优化算法的 k-means 的特征分类

2.1 蝗虫优化算法

k-means 算法是典型的基于原型的目标函数聚类方法,但受初始类聚类中心点选取的影响较

大,直接聚类结果稳定性差,因此,本文采用蝗虫优化算法选取初始聚类中心后再进行分类,所得分类结果较好。GOA 是根据蝗虫在生态环境中的种群迁移和觅食过程进行模拟的一种改进算法^[7-9],其数学模型可表示:

$$X_i = S_i + G_i - A_i X \quad (1)$$

式中 X_i 为第 i 只蝗虫在蝗群中的位置; S_i 为蝗群中其他蝗虫对第 i 只蝗虫的力; G_i 重力对第 i 只蝗虫施加的力; A_i 为环境风力对第 i 只蝗虫的力。则式(1)进一步改为:

$$X_i = r_1 S_i + r_2 G_i + r_3 A_i \quad (2)$$

式中 r_1, r_2, r_3 为 $[0, 1]$ 间的随机数。

$$S_i = \sum_{j=2}^N s(d_{ij}) \hat{d}_{ij} \quad (3)$$

式中 $d_{ij} = |x_i - x_j|$ 表示第 i 只与第 j 只蝗虫之间的距离; $\hat{d}_{ij} = \frac{x_i - x_j}{d_{ij}}$ 表示为距离的单位向量; $s(\delta)$ 为其他蝗虫对某只蝗虫的互动力,其形式由吸引强度 f 和吸引尺度 l 决定,参考其他文献,可写为:

$$s(\delta) = f e^{\frac{-\tau}{l}} - e^{-\tau} \quad (4)$$

式中 通常取 f 取 0.5、 l 取 1.5。

令 \hat{e}_g, \hat{e}_w 分别为沿重力方向的单位向量和风向方向的单位向量, u 为漂移常数,则式(2)中 $G_i = -g \hat{e}_g, A_i = -u \hat{e}_w$,改写为:

$$X_i = \sum_{j=2}^N s(|x_i - x_j|) \frac{x_i - x_j}{d_{ij}} - g \hat{e}_g + u \hat{e}_w \quad (5)$$

由于蝗虫的新位置仅由蝗虫的当前位置、目标值位置和其他蝗虫位置共同决定,可进一步写为:

$$X_i = c \left\{ \sum_{j=2}^N c \frac{ub_d - ib_d}{2} s(|x_i^d - x_j^d|) \frac{x_j - x_i}{d_{ij}} \right\} + \hat{T}_d \quad (6)$$

式中 ub_d, ib_d, \hat{T}_d 表示蝗虫在蝗群中的位置系数,分别表示第 i 只蝗虫的第 d 维变量的上界与下界和第 d 维变量的目标位置;为简化计算,定义为线性递减系数 c :

$$c = c_{\max} - t \frac{c_{\max} - c_{\min}}{T_{\max}} \quad (7)$$

式中 T_{\max} 为最大迭代次数; t 为当前迭代次数;取 $c_{\max} = 1, c_{\min} = 0.000 1$ 。

2.2 基于 GOA 的 k-means 算法

由于蝗虫算法本身的自适应性,使它在解决优化问题中能够取得较好的效果,因此,本文将 GOA 用于 K-means 初始聚类中心的改进,首先优选目标函数为:

$$fit = \min \sum_{i=1}^M \sum_{x \in E_i} dist(x - e_i) \quad (8)$$

式中 M 为聚类个数, E_i 为样本合集; $x \in E_i$ ($i = 1, 2, 3 \dots M$); e_i 为第 i 类的类中心;因此,可得 GOA-K-means 聚类流程图如下:

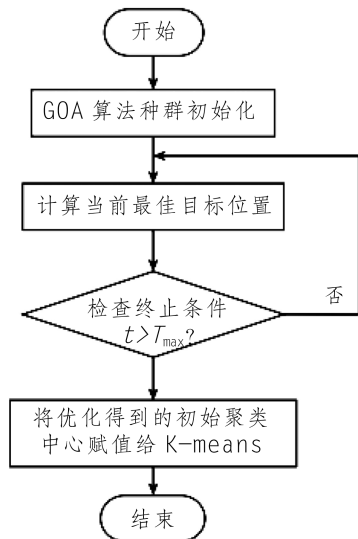


图 1 GOA-K-means 聚类流程图

同时,还可得出如下结果:

- (1) 初始化 GOA 算法参数 T_{max} 、 K 、 D 、 c_{max} 和 c_{min} ;
- (2) 随机生成初始蝗群位置数据;
- (3) 利用式(6)计算蝗虫个体的适应度值,并将最优适应度值的位置作为目标位置;
- (4) 利用式(4)更新位置;
- (5) 根据迭代次数进行迭代;
- (6) 将 GOA 优化得到的最小值赋值给 K-means;
- (7) 根据波形持续时间判断故障进程;
- (8) 结合式(6)中特征完成故障聚类。

3 实际验证

3.1 数据验证

采集贵州某市 2021 年 4~5 月城市配网的录波数据,共记录了发生的 750 次故障及扰动。本次采样频率为 1 200 Hz,每个波形共包含录

波装置触发前的 4 个周波与故障后的 8 个周波。选择其中 170 例作为测试集,其余为训练集和验证集,训练时采用交叉验证。各个集合中的样本个数如下:

表 1 各个集合中的样本个数

故障类型	单周波 早期故障	多周波 早期故障	永久性 故障	暂态 干扰
测试集中数量	71	59	5	35
训练集和验证 集中数量	206	157	6	211
总数	277	216	11	246

利用本文方法进行早期故障分类结果示意

如下:

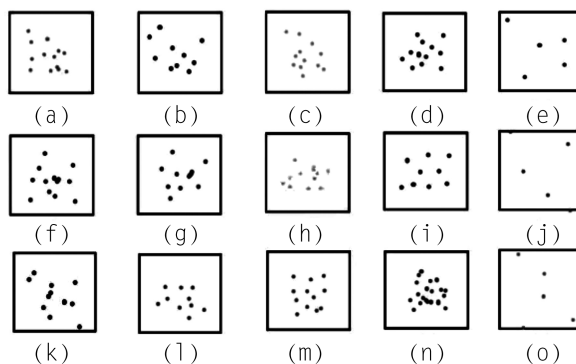


图 2 早期故障分类结果

由于不同位置、不同类型的早期故障表现的暂态波形不一致,因此,本文分类得到了 15 种故障。其中(a-f)为单周波故障共 6 类,(g-l)为多周波故障共 6 类,(m,n)为暂态干扰,(o)为永久性故障。其中单周波故障及多周波故障占全部故障的大多数。又由于某一元件发生的单周波故障和多周波故障在暂态波形上具有一定相似性,本文利用余弦相似度进一步进行分类,可得到同一元件故障波形分类(图 3):

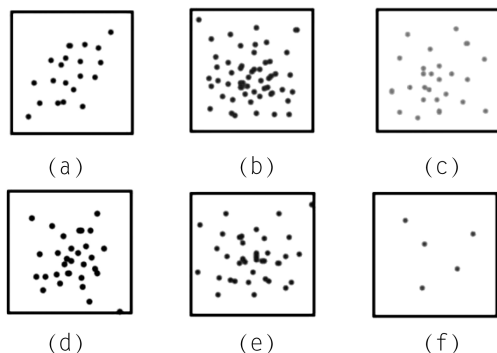


图 3 同一元件故障波形分类图

查看标签,发现其中(e)、(f)分别为暂态干扰和永久性故障,与实际情况一致。其余 4 类中分别含有单周波和多周波早期故障若干,则结论为本文采用的数据中共包含了配网线路中的 4 种不同元件或不同位置的故障,与理论分析一致。按每一类中单周波和多周波的发生时间标签将早期故障进行排列,可得某一故障点的早期故障序列(图 4)。利用该序列可预计每一类故障发展成

永久性故障的发生时间。

可以看出同一类故障之间的单周波与多周波有顺序发展的趋势,各次故障波形之间的时间间隔越来越短,当单周波转变为多周波故障后间隔时间依然存在越来越短的趋势,且故障周波越来越多。因此,可以结合时间数据预计该类故障距离发展成永久性故障的时间。

3.2 评价

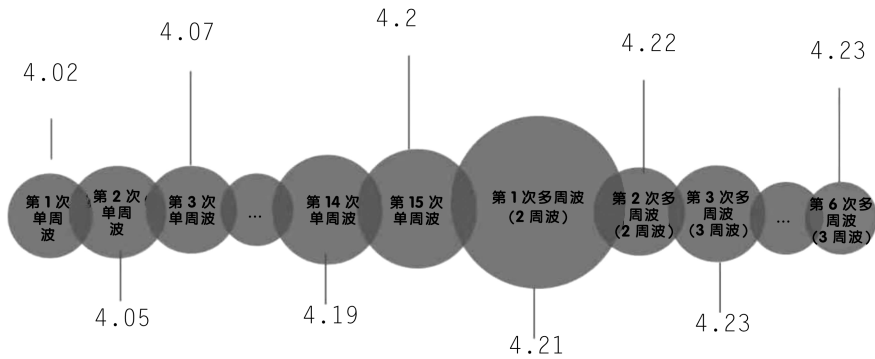


图 4 某一故障点的早期故障序列

设定评价函数 A_{cc} 为准确率、 P_{re} 为精确率、 R_{ec} 为召回率,计算方法如下:

$$A_{cc} = \frac{TP + TN}{TP + TN + FP + FN} \quad (9)$$

$$P_{re} = \frac{TP}{TP + FP} \quad (10)$$

$$R_{ec} = \frac{TP}{TP + FN} \quad (11)$$

式中 TP 为真阳性因子,表示预测为某类早期故障,实际也为该类早期故障; FP 为假阳性因子,表示预测为某类早期故障,实际不为该类早期故障; FN 为假阴性因子,表示预测不为某类早期故障,实际为该类早期故障; TN 为真阴性因子,表示预测不为某类早期故障,实际也不为配网早期故障。

分别求取每一个分类下的评价指标,对比基于小波的早期故障检测方法、基于贝叶斯网络的早期故障检测方法以及卷积神经网络模型,其评价函数值如表所示,可以看出本文使用的方法优于其他方法。不同方法检测结果对比见表 2。

4 结 语

电网早期故障诊断是永久性故障的前期过程,对其进行诊断可以在较宽的时间维度中进行

表 2 不同方法检测结果对比

检测方法	基于小波的早期故障检测方法	基于贝叶斯网络的早期故障检测方法	卷积神经网络模型	本文方法
A_{cc}	0.675	0.868	0.853	0.916
P_{re}	0.712	0.901	0.819	0.912
R_{ec}	0.695	0.874	0.872	0.901

故障排查,但早期故障发生的概率低、信号较弱、持续时间短,影响因素多,诊断困难。本文通过 GOA-k-mean 算法对电缆早期故障进行诊断,通过 GOA 解决 k-means 算法的初始化问题,进而对采集的早期故障的波形进行分类,利用波形的时序特征进行早期故障种类的确定,利用余弦相似度进行故障点的分类,达到找出故障点个数和故障阶段的目的,为精确预测即将发生的故障时间提供基础。本文方法的提出为永久性故障预测提供了一种可行方案,后期将结合长短时神经网络对永久性故障发生时间进行精准预测。

参考文献:

- [1] 汪颖,孙建风,肖先勇,等.基于优化卷积神经网络的电缆早期故障分类识别[J].电力系统保护与控制,2020,48(7):10-18.
- [2] Ying Wang, Hong Lu, Xianyong Xiao, Xiaomei Yang. Cable incipient fault identification using restricted Boltzmann ma-

chine and stacked autoencoder[J]. IET Generation, Transmission & Distribution, 2020, (7).

- [3] 熊思衡, 刘亚东, 方健, 等. 配电网早期故障辨识方法[J]. 高电压技术, 2020, 46(11): 3970-3976.
- [4] 王文凯, 邓斌. 基于 DAE-IPSO-SVM 的电缆早期故障识别方法[J]. 国外电子测量技术, 2021, 40(8): 29-35.
- [5] 侯鹏飞, 马宏忠, 吴金利, 等. 基于混沌理论与蝗虫优化 K-means 聚类算法的电抗器铁芯和绕组松动状态监测[J]. 电力自动化设备, 2020(11): 181-189.
- [6] 徐新, 蒋波涛, 曹雯. 蝗虫优化神经网络在变压器故障诊断中的应用[J]. 电网与清洁能源, 2021, 37(5): 17-23.
- [7] 刘亮, 何庆. 基于改进蝗虫优化算法的特征选择方法[J]. 南京大学学报(自然科学), 2020, 56(1): 41-50.
- [8] KULKARNI S, SANTOSO S. Time-domain algorithm

for locating evolving faults [J]. IEEE Transactions on Smart Grid, 2012, 3(4): 1584-1593.

- [9] DU Y, LIU Y, SHAO Q, et al. Single line-to-ground faulted line detection of distribution systems with resonant grounding based on featurefusion framework [J]. IEEE Transactions on Power Delivery, 2019, 34(4): 1766-1775.

作者简介:

唐辰旭(1988-),男,四川成都人,工程师,工学硕士,主要从事电力系统保护与控制、配网调控、配网运营管控等;

吉畅(1991-),女,贵州毕节人,工程师,工学硕士,主要从事电力系统参数辨识、故障诊断等.

(责任编辑:卓政昌)

(上接第 91 页)

合温度因子、时变因子共计 12 个参数作为大坝变形预测的因子集;

(2)针对传统长周期安全监测数据缺失的问题,基于滑动平均窗口法对缺失数据进行处理,取得了良好效果;

(3)应用随机森林算法对历史变形监测数据进行学习训练,并与常用的数据挖掘方法对比,本文所提出的方法预测精度最高,具有十分突出的优越性;

(4)下一步,可将本文模型应用于大坝安全管理系统中,通过模型预测值与实测值进行对比分析,及时掌握高拱坝服役性态,实现对大坝的精细化管理。

参考文献:

- [1] 张博庭. 中国水电 70 年发展综述——庆祝中华人民共和国成立 70 周年[J]. 水电与抽水蓄能, 2019, 5(5): 1-6.
- [2] 程春田. 碳中和下的水电角色重塑及其关键问题[J/OL]. 电力系统自动化: 1-10. <http://kns.cnki.net/kcms/detail/32.1180.TP.20210203.1442.006.html>.
- [3] 国家能源局大坝安全监察中心. 大坝概览[EB/OL]. [2021. 05. 11]. <https://www.dam.com.cn/overview/dams>.
- [4] 曹延明, 井德泉, 刘春高. 人工免疫算法优化双支持向量机在拱坝变形预测中的应用[J]. 长江科学院院报, 2019, 36(12): 54-58.
- [5] 魏博文, 彭圣军, 徐镇凯, 等. 顾及大坝位移残差序列混沌效应的 GA-BP 预测模型[J]. 中国科学: 技术科学, 2015, 45(5): 541-546.
- [6] 吴中如. 水工建筑物安全监控理论及其应用[M]. 水工建

筑物安全监控理论及其应用, 2003.

- [7] 刘敬洋, 刘何稚, 朱凯, 等. 基于 PSO-SVM 模型的拱坝坝体变形预测研究[J]. 三峡大学学报(自然科学版), 2013, 35(1): 30-33.
- [8] 胡波, 刘观标, 吴中如. 小湾特高拱坝首蓄期坝体变形特性分析及评价[J]. 水利水电科技进展, 2015, 35(6): 68-72.
- [9] 王晓玲, 谢怀宇, 王佳俊, 等. 基于 Bootstrap 和 ICS-MKELM 算法的大坝变形预测[J]. 水力发电学报, 2020, 39(3): 106-120.
- [10] 胡江, 王春红, 马福恒. 特高拱坝运行初期变形预测模型温度因子选取方法[J]. 长江科学院院报, 2021, 38(01): 59-65.
- [11] 李润鸣, 包腾飞, 高瑾瑾, 等. 基于小波 EGM-ISFLA-SVR 的大坝变形组合预测模型[J]. 水利水电技术, 2018, 49(5): 57-62.
- [12] Kang F, Li J. Displacement Model for Concrete Dam Safety Monitoring via Gaussian Process Regression Considering Extreme Air Temperature[J]. Journal of Structural Engineering, 2020, 146(1): 5019001.
- [13] 许江, 陈志奎, 张清辰. 基于嵌套滑动窗口的数据流缺失数据填充算法[J]. 西南师范大学学报(自然科学版), 2015, 40(11): 130-136.
- [14] 方匡南, 吴见彬, 朱建平, 等. 随机森林方法研究综述[J]. 统计与信息论坛, 2011, 26(3): 32-38.

作者简介:

伍远朋(1988-),男,四川成都人,工程师,硕士学位,主要从事水利水电工程建设管理;

刘宗显(1995-),男,陕西安康人,工程师,硕士学位,主要从事水利水电工程建设管理.

(责任编辑:卓政昌)