# 基于均值漂移聚类的开关柜局部放电异常检测

黎阳羊',胡金磊',赖俊驹',王伟',杨帆<sup>2</sup>

(1. 广东电网有限责任公司清远供电局,广东 清远 511500;

2. 上海电力大学 电气工程学院,上海 200090)

摘要:针对开关柜现场带电检测数据,提出一种基于多维特征量的均值漂移聚类算法,对开关柜的局部放 电进行异常识别。采用局部放电检测数据的离散度、平均距离百分比、集中度和最大波动率指标全面量化开 关柜局部放电的程度,并构建多维特征数据库;通过自动搜索偏移量的均值漂移聚类算法对开关柜的状态进 行划分,并通过所给定的开关柜簇标签隶属度函数判定是否为异常点,由此实现开关柜的绝缘状态异常检测。 对现场带电检测实际数据进行实例分析,验证该方法的可行性,为开关柜的绝缘状态异常识别提供一定的理 论依据。

关键词:开关柜;绝缘状态;均值漂移聚类算法;多维特征数据库;异常检测 中图分类号:TM28 文献标识码:A DOI:10.19457/j.1001-2095.dqcd22659

Anomaly Identification of Switchgear Insulation Condition Based on Mean-shift Clustering

LI Yangyang<sup>1</sup>, HU Jinlei<sup>1</sup>, LAI Junju<sup>1</sup>, WANG Wei<sup>1</sup>, YANG Fan<sup>2</sup> (1. Qingyuan Power Supply Bureau of Guangdong Power Grid Inc., Qingyuan 511500, Guangdong, China; 2. School of Electrical Engineering, Shanghai University of Electrical Power, Shanghai 200090, China)

**Abstract:** Aiming at the on-site live detection data of switchgear, a mean-shift clustering algorithm based on multi-dimensional feature quantity was proposed to identify the abnormal partial discharge of switchgear. The switchgear partial discharge detection data which include dispersion, average distance percentage, concentration and maximum volatility were used to quantify the insulation of switchgear to construct multi-dimensional feature database. The insulation condition of the switchgear was divided by mean-shift clustering algorithm which can automatically search for the offset, and the abnormal point was determined by the membership function of the cluster labels to realize the abnormal detection of the switchgear insulation condition. The feasibility of the algorithm was verified by the live detection data, which can provide a certain theoretical basis for the switchgear evaluation of the switchgear.

Key words: switchgear; insulation condition; mean-shift clustering algorithm; multi-dimensional feature database; anomaly identification

随着我国电网的不断发展,电力系统对电力 设备的运行可靠性要求也越来越高,对电力设备 的状态检测、评价与运维有着更高的要求<sup>[1-3]</sup>。目 前,开关柜的评价方法主要有两种,一种是基于 模糊综合、D-S论据(dempster-shafer evidential theory)等状态评价方法<sup>[4-6]</sup>,适用于开关柜停电检 修时对开关柜进行全面细致的评估,包括了从目 测柜体状态、表记状态、内部部件测试等方面提 取特征量,在实际检修中停电检修的周期较长; 另一种是通过聚类算法等机器学习算法在开关 柜常规的带电检修中,对开关柜的局部放电程度 进行评估<sup>[7-8]</sup>。该方法主要适用于开关柜的带电 检修,实际检修中带电检修的周期远小于停电检 修周期。目前现有文献普遍未对开关柜局部放 电的检测机理进行全面的分析和研究,缺乏从检 测机理入手构建开关柜的局部放电多维特征量。

开关柜的带电检测方法包括暂态对地电压 (transient earth voltages, TEV)检测法和超声检测

**基金项目:**广东电网有限责任公司科技项目(031800KK52170056) 作者简介:黎阳羊(1990一),男,本科,Email:744696264@qq.com

法,TEV的检测原理是检测放电点从柜体金属表面的断开处或者缝隙处随机散射的电磁信号,超声检测是检测局部放电活动中的声波辐射信号<sup>[9]</sup>。IEEE的指导准则中指出由于型号不一致、结构复杂、背景噪声以及不同检测传感器厂商的量化信号方式不同等原因,缺乏对开关柜绝缘状态进行全面的数据信息挖掘分析,单一的柜体检测难以对评估开关柜的局部放电的大小、劣化程度做出合理量化<sup>[10]</sup>。本文提出了一种全面的开关柜局部放电程度的状态评估技术。根据日常巡检中柜体前、后面的上、中、下6个检测点的TEV以及超声数据<sup>[11]</sup>,结合机理分析,引入检测数据的离散度、平均距离百分比、集中度以及最大波动率等指标量化开关柜的局部放电程度。

此外,本文对日常对各地区的各开闭所、环 网柜以及分支箱的数据进行调研,局部放电异常 的样本数据多为小规模离群异常数据。如文献 [7]中某变电站绝缘状态异常的开关柜仅占总数 的6.82%。文中所提出的基于距离的均值聚类算 法,当异常点偏移正常样本较大且比例较少,其 鲁棒性较差。而均值漂移聚类是基于滑动窗口 的算法,其目的找到绝缘状态特征量所处的密集 区域,不受异常点的影响。由于每次滑动的窗口 随着样本点的变化,其偏移量对均值偏移向量的 贡献也不同。因此,本文通过高斯核函数自动量 化均值漂移方向与大小,并基于文献[12-13]对绝 缘状态进行划分。同时为了自动判别聚类算法 所划分的簇是否为异常点,本文给定簇标签隶属 度函数判断该簇是否为异常点<sup>[14]</sup>。

综上,本文依据柜体和环境参数及局部放电 检测数据的离散度、平均距离百分比、集中度以 及最大波动率4个指标全面量化开关柜绝缘劣化 状态,同时建立多维特征库。采用基于高斯核函 数的自动寻参均值漂移聚类算法对绝缘状态进 行划分,并给定簇标签隶属度函数自动判断该簇 是否为异常点,由此实现开关柜的异常检测。

# 1 开关柜局部放电的波动性与局部 放电异常检测流程

参考国家电网公司的 Q/GDW645—2011 标 准《配网设备状态评价导则》,从柜体 TEV 和超声 局部放电检测评估开关柜的局部放电程度。局 部放电的检测机理如图 1 所示,假设 D 为放电源, 即当 D 处产生局部放电时,会产生散射电磁波信 号的现象,TEV 主要检测其从开关柜缝隙发出的 电磁波,而开关柜的柜体是密闭的金属外壳,电 磁信号无法穿透,因此测量局部放电异常的随机 性比较高。结合如表1所示的日常各地区的各开 闭所、环网柜以及分支箱的带电检测样本报告统 计中典型的带电检测局放异常数据(通常以对数 单位"dB"来表示放电强度)。异常数据具有较强 的波动性——主要体现在整体偏高或局部偏高。 对于超声信号,同样存在漫散射问题。



表1 典型的带电检测局放异常数据

Tab.1 Classical online detection abnormal data

护旦	测具店	兆垦店/JD	测量占	放电强度/dB		
姍与	侧里阻	月 泉 沮/ɑb	侧里品	上	中	下
	(DD) 7	5	柜体正面	4	4	5
G05 PT柜 超声	3	柜体反面	3	4	5	
	超声	6	柜体正面	7	6	42
			柜体反面	6	8	33
G116 PT柜	TEV	2	柜体正面	14	16	20
	ILV	3	柜体反面	15	16	11
	切吉	4	柜体正面	4	5	4
	旭戸	超声 4	柜体反面	4	4	5

本文提出检测柜体前、后面的上、中、下共计 6个检测点的离散度、集中度以及最大波动率量 化开关柜的局部放电异常的情况。其中,通过离 散度能反映开关柜绝缘状态数据的离散情况;平 均距离百分比反映数据的偏移情况;集中度反映 样本所处的中间水平;最大波动率反映样本数据 的离差情况。

对多维度的开关柜特征量进行绝缘状态异 常检测时,开关柜的原始绝缘状态样本数据劣化 程度极少,通过全面的波动性量化将提高均值漂 移聚类对异常点识别的鲁棒性。在均值漂移聚 类算法中,首先可以确定处于密度大的样本点所 处的类别,即聚类的簇,再对异常点进行逐步搜 索。开关柜绝缘状态数据集的形状是未知的,而 均值漂移聚类可以对任意形状的数据集进行聚 类。K均值聚类算法对小规模的异常样本点识别 能力较弱。

如图2所示,均值漂移聚类算法选取初始中 心点与半径为R的滑动窗口朝向密度大的聚类点 进行漂移聚类,每次窗口滑动的方向和大小为向 量*M*<sub>h</sub>,初始点加*M*<sub>h</sub>得出下一中心点,对其进行迭 代计算,直至滑动窗口的偏移量收敛。该算法对 沿密度最大的方向自动搜索,不需要人为设定算 法的聚类个数。此外,均值漂移聚类结果没有偏 倚,由于是遍历整个样本点并进行以半径为R的 圆形自动搜索,初始聚类中心的位置对聚类结果 的影响较小。相对的,K均值聚类算法对初始值 非常敏感。



图 2 均值漂移聚类示意图 Fig.2 Diagram of mean-shift cluster

2 基于波动性指标的开关柜局放多 维特征量数据库

#### 2.1 开关柜绝缘状态波动性指标

2.1.1 减少开关柜背景值干扰

开关柜的带电检测数据包括6个检测点的 TEV、超声的幅值检测量以及TEV和超声的背景 值。首先计算TEV以及超声的幅值偏差量:

$$T_i^{(k)} = (To_i^{(k)} - T_B^{(k)}) / To_i^{(k)}$$
(1)

$$S_i^{(k)} = (So_i^{(k)} - S_B^{(k)})/So_i^{(k)}$$
(2)

式中: $To_i^{(k)}$ 为TEV的幅值第 $i(i=1,2,\dots,6)$ 个检测 点的第 $k(k=1,2,\dots,n)$ 台开关柜的检测值; $T_B^{(k)}$ 为 TEV的背景值; $T_i^{(k)}$ 为TEV的幅值偏差量; $So_i^{(k)}$ 为 超声的幅值检测值; $S_B^{(k)}$ 为超声的背景值; $S_i^{(k)}$ 为超 声的幅值偏差量。

2.1.2 偏差量的惩罚函数

为了克服数据中含有负数或者零所带来的 数据爆炸的问题——最大波动率不能为零,本文 对偏差量引入惩罚项,使得整体的数据全部大于 零,计算公式如下:

$$T_i^{(k)\prime} = T_i^{(k)} + |\text{INT}(\min\{T_1^{(k)}, T_2^{(k)}, \cdots, T_6^{(k)}\})| \quad (3)$$

$$S_i^{(k)\prime} = S_i^{(k)} + |\text{INT}(\min\{S_1^{(k)}, S_2^{(k)}, \cdots, S_6^{(k)}\})| \quad (4)$$

式中:INT(min { $T_1^{(k)}, T_2^{(k)}, \dots, T_6^{(k)}$ })为当前TEV 偏差 量的最小值的取整函数;INT(min { $S_1^{(k)}, S_2^{(k)}, \dots, S_6^{(k)}$ }) 为当前超声偏差量的最小值的取整函数。

通过计算出的 $T_i^{(k)'}$ 和 $S_i^{(k)'}$ 更新幅值偏差量 $T_i^{(k)}$ 和 $S_i^{(k)}$ 。

2.1.3 离散度

离散度能够反映开关柜局部放电检测数据 的波动情况,分别对6个检测点的TEV和超声检 测偏差量求离散程度。TEV的离散度如下式所示:

$$D_{\rm T}^{(k)} = \sqrt{\frac{1}{m-1} \sum_{i=1}^{m} (T_i^{(k)} - \overline{T^{(k)}})}$$
(5)

式中:m为数据点的数量; $T_i^{(k)}$ 为TEV幅值偏差量;  $\overline{T^{(k)}}$ 为该台开关柜所有检测点数据偏差量的均值;  $D_T^{(k)}$ 为开关柜TEV检测数据偏差量的离散度。 同理,超声的离散度计算式如下式:

$$D_{\rm S}^{(k)} = \sqrt{\frac{1}{m-1} \sum_{i=1}^{m} (S_i^{(k)} - \overline{S^{(k)}})}$$
(6)

式中:*S*<sup>(k)</sup>为超声幅值偏差量;*S*<sup>(k)</sup>为该台开关柜所 有检测点数据偏差量的均值;*D*<sup>(k)</sup>为开关柜超声 检测数据偏差量的离散度。

2.1.4 平均距离百分比

平均距离百分比为该台开关柜所有检测数 据距离中心点的距离与均值之比,通过平均距离 百分比揭示开关柜绝缘检测数据整体变动的程 度,如下式所示:

$$A_{\rm T}^{(k)} = \sqrt{\frac{1}{\overline{T^{(k)}}} \frac{1}{n} \sum_{i=1}^{n} |T_i^{(k)} - \overline{T^{(k)}}|}$$
(7)

$$A_{\rm S}^{(k)} = \sqrt{\frac{1}{\overline{S^{(k)}}} \frac{1}{n} \sum_{i=1}^{n} |S_i^{(k)} - \overline{S^{(k)}}|} \tag{8}$$

式中:A<sub>T</sub><sup>(h)</sup>为开关柜TEV 检测数据偏差量的平均 距离百分比;A<sub>S</sub><sup>(h)</sup>为开关柜超声检测数据偏差量 的平均距离百分比。

2.1.5 集中度

集中度可以反映样本中数据的集中情况,通 过6个测量的中数与众数之和的均值反映样本 的整体情况。众数主要是对重复出现的数据进 行记录;中数则与数据的排列位置有关,中数不 受极大或极小值影响,通过观察中数可以找出 开关柜的整体局部放电偏高的情况,当集中度 越高时,局部放电状态越严重。其计算公式如 下式:

$$F_{\rm T}^{(k)} = \frac{Med \{T_i^{(k)}|i=1,2,\cdots,n\} + Mo \{T_i^{(k)}|i=1,2,\cdots,n\}}{2}$$

(9) 65

$$F_{\rm S}^{(k)} = \frac{Med \{S_i^{(k)}|i=1,2,\cdots,n\} + Mo\{S_i^{(k)}|i=1,2,\cdots,n\}}{2}$$
(10)

式中: $Med \{T_i^{(k)}|i = 1,2,...,n\}$ 为该台开关柜TEV幅 值的中值; $Mo \{T_i^{(k)}|i = 1,2,...,n\}$ 为该台开关柜 TEV幅值的中位数; $Med \{S_i^{(k)}|i = 1,2,...,n\}$ 为该台 开关柜超声幅值的中值; $Mo \{S_i^{(k)}|i = 1,2,...,n\}$ 为 该台开关柜超声幅值的中位数。

2.1.6 最大波动率

最大波动率能够量化开关柜局放程度的最 大波动情况,区分数据的离差程度,反映数据区 间跨度,如下式所示:

$$M_{\rm T}^{(k)} = \frac{\max \left| T_i^{(k)} - T_j^{(k)} \right|}{\overline{T^{(k)}}} \tag{11}$$

$$M_{\rm s}^{(k)} = \frac{\max \left| S_i^{(k)} - S_j^{(k)} \right|}{\overline{S^{(k)}}} \tag{12}$$

式中:T<sub>i</sub><sup>(k)</sup>,T<sub>j</sub><sup>(k)</sup>为第k台开关柜任意两个测量点的 TEV幅值偏差量;M<sub>T</sub><sup>(k)</sup>为开关柜TEV检测数据偏 差量的最大波动率;S<sub>i</sub><sup>(k)</sup>,S<sub>j</sub><sup>(k)</sup>为第k台开关柜任意 两个测量点的超声幅值偏差量;M<sub>s</sub><sup>(k)</sup>为开关柜超 声检测数据偏差量的最大波动率。

#### 2.2 多维特征数据库

依据式(5)~式(12)的各台开关柜特征量的 计算结果,开关柜的多维特征数据库可建立为

$$\boldsymbol{R} = \begin{bmatrix} D_{\mathrm{T}} & D_{\mathrm{S}} & A_{\mathrm{T}} & A_{\mathrm{S}} & F_{\mathrm{T}} & F_{\mathrm{S}} & M_{\mathrm{T}} & M_{\mathrm{S}} \end{bmatrix}$$
$$= \begin{bmatrix} r_{1,1} & r_{1,2} & \cdots & r_{1,8} \\ r_{2,1} & r_{2,2} & \cdots & r_{2,8} \\ \vdots & \vdots & \ddots & \vdots \\ r_{k,1} & r_{k,2} & \cdots & r_{k,8} \end{bmatrix} = \begin{bmatrix} r_{1} \\ r_{2} \\ \vdots \\ r_{k} \end{bmatrix}$$
(13)

式中:**R**为开关柜的多维特征数据库;**r**<sub>i</sub>(*i*=1,2, …,*k*)为各台开关柜的特征量,例如*r*<sub>k8</sub>代表第*k*个开关柜的第8个特征。

对所处理的特征数据进行 z-score 标准化, 消除数据数量级之间的差异。

## 3 基于均值漂移聚类的绝缘状态异 常检测法

均值漂移聚类是基于滑动窗口的算法,其目的是找到绝缘状态特征量所处的密集区域。它的目标是定位每类的中心点,通过将中心点的 候选点更新为滑动窗口内点的均值来完成。然 后消除近似重复,形成最终的中心点集及其相应 的组。

对于给定的d维度空间 $\mathbf{R}^{d}$ 中的样本点 $\mathbf{r}_{i}(i=1, 66)$ 

2,…,k),半径为 $S_h$ 的高维球区域,其mean-shift向量的基本形式为

$$\boldsymbol{M}_{\rm h} = \frac{1}{k} \sum_{\boldsymbol{r}_i \in S_{\rm h}} (\boldsymbol{r}_i - \boldsymbol{r}) \tag{14}$$

其中,**r**为任意选取的初始算法样本中心点,则在 第*t*个中心点的漂移为

$$\boldsymbol{r}^{\iota+1} = \boldsymbol{M}_{\rm h}^{\iota} + \boldsymbol{r}^{\iota} \tag{15}$$

遍历高维球区域内所有的向量,求出 meanshift向量,此时该向量称为均值漂移向量。因此, 对于 mean-shift 均值漂移算法,通过算出当前点 的漂移均值并进行迭代计算,并以漂移均值为新 的起始点,直到其收敛为止。

在 mean-shift 算法中引入核函数使得随着样本与被偏移点的距离不同,其偏移量对均值偏移向量的贡献也不同。因此引入高斯核函数量化偏移度,点r的核密度函数估计为

$$f(\mathbf{r}) = \frac{C_k}{nh^d} \sum_{i=1}^n K ||\frac{\mathbf{r} - \mathbf{r}_i}{h}||^2$$
(16)

式中:*C<sub>k</sub>*为常量;*K*为核函数;*h*为核宽。 对式(16)求导,寻找局部密度极大点:

$$\hat{f}(\mathbf{r}) = \frac{2C_k}{nh^{d+2}} \sum_{i=1}^n (\mathbf{r} - \mathbf{r}_i) \cdot K' \left( ||\frac{\mathbf{r} - \mathbf{r}_i}{h}||^2 \right) \quad (17)$$

令 $g(\mathbf{r}) = -K'(\mathbf{r})$ ,式(17)可以表示为

$$\hat{f}(\mathbf{r}) = \frac{2C_k}{nh^{d+2}} \sum_{i=1}^n g(||\frac{\mathbf{r} - \mathbf{r}_i}{h}||^2) \cdot \mathbf{M}_{\rm h}(\mathbf{r})$$
(18)

其中

$$\boldsymbol{M}_{h}(\boldsymbol{r}) = \left[\sum_{i=1}^{n} \boldsymbol{r}_{i} g\left(||\frac{\boldsymbol{r}-\boldsymbol{r}_{i}}{h}||^{2}\right)\right] \cdot \left[\sum_{i=1}^{n} g\left(||\frac{\boldsymbol{r}-\boldsymbol{r}_{i}}{h}||^{2}\right)\right]^{-1} - \boldsymbol{r}$$
(19)

式(19)为基于高斯核函数的均值漂移向量。

在采用均值漂移聚类算法时,其聚类后的结 果呈现出绝缘状态劣化程度差的时候样本点极 少,个别情况下,单一的样本点是一个独立的簇。 而处于平稳运行下的开关柜数据极多。

依据开关柜绝缘状态标签从优秀到异常会 急剧减少的特性,本文采用如图3所示的簇标签



Fig.3 Membership function of cluster label

隶属度函数,隶属度函数为等间隔函数,并依据 等间隔制定从优秀到异常5个标签集。通过隶属 度函数,找出样本点所在簇的标签。簇内样本点 密度定义为 D(k) = 1 - N(k)/N (20)

式中:D(k)为第k个簇的样本密度;N(k)为第k个 簇的样本点个数;N为总样本点个数。 综上,本文流程图如图4所示。

际上, 平义流住图如图4 所小



图 4 开关柜绝缘状态异常检测流程图 Fig.4 Switchgear insulation condition anomaly identification flow char

### 4 案例分析

选用某地变电站现场设备检测报告共计299 组10kV高压开关柜带电检测实测数据作为数据 集,对上述算例采用实例分析。带电检测数据包 括开关柜前、后面的上、中、下6个检测点的TEV 和超声检测数据、背景噪声数据。案例分析主要 包括了两方面:

1)本文基于所提出的波动性与均值漂移聚 类,把未经过波动性处理与均值漂移聚类相对 比,分别与专家根据《配网设备状态评价导则》对 开关柜局部放电的评估结果对比,计算局部放电 异常检测的准确率;

2)波动性处理后的K均值聚类算法与均值漂 移聚类算法相对比,与专家根据《配网设备状态 评价导则》对开关柜局部放电的评估结果对比, 计算局部放电异常检测的准确率。

#### 4.1 绝缘状态异常检测结果

应用局部放电多维特征量的构建方法,建立 如式(13)的开关柜绝缘状态多维样本特征数据 库 R,并对数据库进行标准化预处理。通过均值 漂移聚类算法,对开关柜的绝缘状态进行均值漂 移聚类划分,聚类中心点如表2所示,各簇之间中 心点划分明确。

根据图2所示的簇标签隶属度函数计算所得 的各类别标签如表3所示。

表2 基于均值漂移聚类算法的开关柜绝缘状态聚类中心点

Tab.2	with mean-	shift clustering	algorithm	er centers
5 見引	D			4

类别	$D_{\mathrm{T}}$	Ds	$A_{\mathrm{T}}$	As
1	-0.123	0.181	-0.360	-0.345
2	-0.469	0.197	-0.551	-0.674
3	4.634	1.394	1.483	1.882
4	-0.251	-0.229	-0.323	-0.359
5	8.815	5.243	1.008	1.514
6	4.568	0.294	2.912	2.917
7	-0.304	4.060	-0.854	-1.090
8	-0.304	0.019	-0.620	-0.446
9	-0.556	-1.081	-0.201	-0.203
类别	$F_{\mathrm{T}}$	$F_{\rm S}$	$M_{\mathrm{T}}$	M <sub>s</sub>
类别 1	<i>F</i> <sub>т</sub> -0.167	F <sub>S</sub> -0.079	М <sub>т</sub> -0.327	M <sub>s</sub> -0.322
类别 1 2	F <sub>T</sub> -0.167 2.025	<i>F</i> <sub>s</sub> -0.079 6.746	М <sub>т</sub> -0.327 -0.622	<i>M</i> <sub>s</sub> -0.322 -0.844
<u>类别</u> 1 2 3	F <sub>T</sub> -0.167 2.025 0.288	$F_{\rm S}$ -0.079 6.746 -0.309	$M_{\rm T}$ -0.327 -0.622 0.410	$M_{\rm S}$ -0.322 -0.844 0.670
<u>类别</u> 1 2 3 4	$F_{\rm T} = -0.167$ 2.025 0.288 6.973	$F_{\rm S} = -0.079$ 6.746 -0.309 4.655	$\frac{M_{\rm T}}{-0.327}$ -0.622 0.410 0.593	$\frac{M_{\rm S}}{-0.322}$ -0.844 0.670 0.976
<u>类别</u> 1 2 3 4 5	$F_{\rm T} = -0.167$ 2.025 0.288 6.973 -0.040	$\begin{array}{c} F_{\rm S} \\ -0.079 \\ 6.746 \\ -0.309 \\ 4.655 \\ -0.416 \end{array}$	$\begin{array}{c} M_{\rm T} \\ -0.327 \\ -0.622 \\ 0.410 \\ 0.593 \\ 0.166 \end{array}$	$\begin{array}{c} M_{\rm S} \\ -0.322 \\ -0.844 \\ 0.670 \\ 0.976 \\ 0.352 \end{array}$
类别 1 2 3 4 5 6	$F_{\rm T} = -0.167$ 2.025 0.288 6.973 -0.040 -0.638	$\begin{array}{c} F_{\rm S} \\ \hline -0.079 \\ 6.746 \\ -0.309 \\ 4.655 \\ -0.416 \\ -0.496 \end{array}$	$\begin{array}{c} M_{\rm T} \\ -0.327 \\ -0.622 \\ 0.410 \\ 0.593 \\ 0.166 \\ -1.001 \end{array}$	$\frac{M_{\rm S}}{-0.322}$ -0.844 0.670 0.976 0.352 -2.254
类别 1 2 3 4 5 6 7	$\begin{array}{c} F_{\rm T} \\ -0.167 \\ 2.025 \\ 0.288 \\ 6.973 \\ -0.040 \\ -0.638 \\ -0.040 \end{array}$	$\begin{array}{c} F_{\rm S} \\ -0.079 \\ 6.746 \\ -0.309 \\ 4.655 \\ -0.416 \\ -0.496 \\ -0.656 \end{array}$	$\begin{array}{c} M_{\rm T} \\ -0.327 \\ -0.622 \\ 0.410 \\ 0.593 \\ 0.166 \\ -1.001 \\ 2.501 \end{array}$	$\begin{array}{c} M_{\rm S} \\ -0.322 \\ -0.844 \\ 0.670 \\ 0.976 \\ 0.352 \\ -2.254 \\ 2.259 \end{array}$
类别 1 2 3 4 5 6 7 8	$F_{\rm T} = -0.167$ 2.025 0.288 6.973 -0.040 -0.638 -0.040 6.568	$\begin{array}{c} F_{\rm S} \\ -0.079 \\ 6.746 \\ -0.309 \\ 4.655 \\ -0.416 \\ -0.496 \\ -0.656 \\ 1.184 \end{array}$	$\begin{array}{c} M_{\rm T} \\ -0.327 \\ -0.622 \\ 0.410 \\ 0.593 \\ 0.166 \\ -1.001 \\ 2.501 \\ 0.876 \end{array}$	$\begin{array}{c} M_{\rm s} \\ -0.322 \\ -0.844 \\ 0.670 \\ 0.976 \\ 0.352 \\ -2.254 \\ 2.259 \\ 1.562 \end{array}$

黎阳羊,等:基于均值漂移聚类	的开关柜局部放电异常检测
----------------	--------------

#### 表3 均值漂移聚类算法簇内样本点个数

Tab.3 The number of sample points under differe	ent
---	-----

clusters with mean-shift cluster							
类别	样本个数	标签	类别	样本个数	标签		
1	268	优秀	6	3	异常		
2	5	异常	7	12	异常		
3	6	异常	8	1	异常		
4	2	异常	9	1	异常		
5	1	异常					

类别1的标签个数占样本总数的绝大多数, 通过簇标签隶属度函数可得:所处于该簇的开关 柜的绝缘状态等级为优秀,该簇下的开关柜未有 明显的局部放电现象,可按照正常检测周期进行 带电检测。而类别2~类别9中,绝缘状态的样本 点的个数较少,存在着绝缘状态劣化的现象,建 议运维人员加强关注缩短带电检测周期。

因此,基于表2以及漂移聚类的结果,其绝缘 状态异常检测结果如图5所示。



图5 基于均值漂移聚类的开关柜局部放电异常检测结果

Fig.5 Switchgear insulation condition anomaly identification results with mean-shift cluster

#### 4.2 与其他方法比较的绝缘状态异常检测结果

分别对未经过波动性处理与经过波动性处 理的开关柜带电检测数据进行绝缘状态异常检 测。未经波动性处理的绝缘状态异常检测准确 率为92.64%,处理后的准确率提升到了97.99%, 可见波动性指标能更全面地量化开关柜的局部 放电情况,切实有效地提升开关柜绝缘状态异常 检测的准确率。

采用基于欧式距离的 K 均值聚类对绝缘状态进行异常检测,其不同 K 值下的准确率如表4 所示,当 K=8,此时准确率最高。但由于部分特征量离群度大,且 K 均值聚类对异常点的鲁棒性较差,导致其准确率与均值漂移聚类有一定的差距。其聚类中心点如表5 所示。基于 K 均值 聚类算法的样本中心点划分不够明确,对异常 值较敏感,导致如类别1 与类别3 聚类中心点的 划分不明确,出现中心点偏移的情况。且 K 由于 需要人为提前指定使得判定是否为异常点较为 困难。

表4 K均值聚类在不同簇下的状态评价准确率

Tab.4	State evaluation	accuracy	of $K$ -means
1 a +	State evaluation	accuracy	of K-means

clustering under different clusters					
类别K	准确率/%	类别K	准确率/%		
<i>K</i> =2	78.26	<i>K</i> =6	78.93		
<i>K</i> =3	74.25	<i>K</i> =7	78.93		
<i>K</i> =4	75.59	<i>K</i> =8	79.26		
<i>K</i> =5	78.26	K=9	78.93		

表5 基于*K*均值聚类算法的绝缘状态聚类中心点

Fab.5	The switchgear insulation condition cluster centers
	with K-means clustering algorithm

with K means clustering argorithm						
类别	$D_{\mathrm{T}}$	$D_{\rm S}$	$A_{\mathrm{T}}$	$A_{\rm S}$		
1	-0.311	-0.329	0.834	-0.428		
2	2.674	2.585	0.053	0.357		
3	-0.333	-0.303	-0.417	-0.167		
4	-0.621	-0.844	0.130	-0.421		
5	0.976	1.404	-0.019	0.394		
6	-0.381	-0.374	0.299	1.405		
7	-0.049	-0.059	-0.467	-0.048		
8	-0.312	-0.403	-0.698	-0.725		
类别	$F_{\mathrm{T}}$	$F_{\rm S}$	$M_{ m T}$	$M_{\rm S}$		
1	-0.103	-0.134	0.241	-0.380		
2	-0.429	-0.033	0.137	-0.318		
3	-0.057	-0.267	-1.824	2.309		
4	-0.921	-0.469	0.198	-0.551		
5	0.356	-0.340	-0.380	-0.366		
6	0.266	-0.156	-0.380	-0.405		
7	-1.244	5.007	0.316	1.364		
8	0.311	-0.003	0.602	-0.313		

综上所述,通过波动性处理,能够全面地反 映开关柜绝缘状态。此外开关柜的绝缘状态异 常点离群度较大,通过均值漂移聚类能避免异常 点过大所导致的聚类中心点偏移的情况。

### 5 结论

1)本文所提的波动性指标能够反映对开关 柜整体放电的突变程度。研究结果表明,通过波 动性处理后的绝缘状态特征量能有效提升开关 柜绝缘状态异常检测的准确率。

2)通过基于高斯核函数的自动寻参均值漂 移聚类算法对绝缘状态进行划分,并给定簇标签 隶属度函数自动判断该簇是否为异常点,能有效 实现开关柜的异常检测。仿真案例表明所提方 法对异常点有较好的鲁棒性,提高开关柜绝缘状 态异常检测的准确率。

#### 参考文献

[1] 刘宇轩,赵峰.城轨车辆辅助三相变压器健康状态评估[J].
 电气传动,2020,50(11):89-94.

Liu Yuxuan, Zhao Feng. Health assessment of three-phase transformer aided by urban rail vehicle[J]. Electric Drive, 2020, 50 (11):89–94.

[2] 邹阳,何倩玲,蔡金锭.基于组合赋权-双基点法的变压器油
 纸绝缘状态综合评估[J].电工技术学报,2019,34(20):
 4400-4408.

Zou Yang, He Qianling, Cai Jinding. Comprehensive evaluation of transformer oil-paper state based on combined weight-double base point method[J]. Transactions of China Electrotechnical Society, 2019, 34(20):4400-4408.

 [3] 李建生,梁军,贠志皓,等.基于电气信息评估设备状态渐进 变化的概率分析方法[J].电工技术学报,2013,28(10):355-363.

Li Jiansheng, Liang Jun, Yun Zhihao, *et al.* Probability analysis of progressive changes of equipment state based on electrical information[J]. Transactions of China Electrotechnical Society, 2013,28(10):355–363.

- [4] 谢静,束洪春,王科,等.基于FCM算法的高压开关柜局部放 电状态评价方法研究[J].高压电器,2015,51(10):82-90,96.
  Xie Jing,Shu Hongchun, Wang Ke, *et al.* Study on state evaluation of partial discharge in high voltage switchgear based on fuzzy clustering algorithm[J]. High Voltage Apparatus, 2015,51 (10):82-90,96.
- [5] 谢静,東洪春,王科.基于模糊分层理论的高压开关柜状态 评估算法[J].高电压技术,2014,40(10):3196-3192.
  Xie Jing, Shu Hongchun, Wang Ke. State evaluation based on AHP fuzzy theory of high voltage switchgears[J]. High Voltage Engineering,2014,40(10):3196-3192.
- [6] 贾亚楠,刘东明,随慧斌.基于正态云模型和D-S证据理论的开关柜运行状态综合评估[J].高压电器,2017,53(9): 247-252.

Jia Yanan, Liu Dongming, Sui Huibin, et al. Comprehensive evaluation for switchgear based on the normal cloud model and

[7] 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.

D–S[J]. High Voltage Apparatus , 2017 , 53(9) : 247-252.

Kang Chaoqun, Li Erxia, Sheng Wanxin, *et al.* Dynamic condition assessment for distribution vacuum switch cabinets based on multi-source information fusion[J]. High Voltage Apparatus, 2017,53(3):235-241.

- [8] 谢静,束洪春,王科,等.基于突变级数法的高压开关柜状态 评价算法[J].高电压技术,2014,40(8):2372-2380.
  Xie Jing, Shu Hongchun, Wang Ke, *et al.* State evaluation algorithm based on catastrophe progression method of high voltage switchgears[J]. High Voltage Engineering, 2014,40(8):2372-2380.
- [9] 李德军, 沈威, 郭志强.GIS局部放电常规检测和超声波检测 方法的应用比较[J]. 高压电器, 2009, 45(3):99-103.
  Li D, Shen W, Guo Z. Application comparison between conventional and ultrasonic detection methods for GIS partial discharge
  [J]. High Voltage Apparatus, 2009, 45(3):99-103.
- [10] Mashikian M S. Preventive maintenance testing of shielded power cable systems[J]. IEEE Transactions on Industry Application, 2002, 38(3):736-743.
- [11] 林勇,徐钦炜,张俊强,等.基于ECP的轨道交通直流牵引供 电系统开关状态评价算法[J].电力系统保护与控制,2018,46 (5):16-24.

Lin Yong, Xu Qinwei, Zhang Junqiang, *et al.* Status evaluation algorithm of DC traction power supply system switchgear based on ECP for rail transit[J]. Power System Protection and Control, 2018,46 (5):16–24.

- [12] Enriquez Alex, Lima Shigeaki, Saavedra Osvaldo R. K-NN and mean-shift algorithm applied in fault diagnosis in power transformers by DGA[C]//20th International Conference on Intelligent System Application to Power Systems (ISAP), New Delhi, India, 2019:1–6.
- [13] 崔芮华,胡文达,耿丽恺.基于小波重构信号奇异点的航空 故障电弧检测[J].电气传动,2018,48(6):69-72.
  Cui Ruihua, Hu Wenda, Geng Likai. Arc cault detection method based on singularity of the wavelet reconstruction signal in aviation power system[J]. Electric Drive,2018,48(6):69-72.
- [14] 向楠,张向文.电动汽车再生制动模糊神经网络控制策略研究[J].电气传动,2020,50(7):86-91.

Xiang Nan, Zhang Xiangwen, *et al.* Electric vehicle regenerative braking control strategy of fuzzy neural network research[J]. Electric Drive, 2020, 50(7):86–91.

> 收稿日期:2020-11-09 修改稿日期:2020-12-03