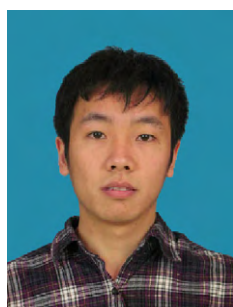


变压器越限特征分群研究

基于变压器越限运行特性，提取越限特征指标体系，包括越限运行总时间、越限运行频次和平均越限运行时间；以广东电网 500 kV 主变压器历史数据为分析对象，对数据进行标准化处理，提取变压器越限特征样本；利用 DBSCAN 算法对变压器越限特征样本集进行聚类分群，分析变压器各分群特征，为掌握大量变压器负载特性提供指导依据。

杜双育¹ 范颖¹ 云伟标² 王路²/1.广东电网有限责任公司电力科学研究院 2.广州泰迪智能科技有限公司



杜双育/博士

关键词/Keywords

变压器 ·
越限运行 ·
DBSCAN算法 ·
聚类 ·

0 引言

随着用电需求量的增加，变压器的高负载甚至越限运行^[1-2]情况也越来越频繁，如统计分析广东电网 500 kV 主变压器平均越限运行时间达 530 min。变压器负载情况特别是越限运行情况，一方面在一定程度上反映了变压器的外部运行环境，如地区负载水平、运行方式等；另一方面更是直观地体现了变压器的绝缘结构老化、寿命损失情况。因此，分析单台变压器的负载特性，并从省、地等区域范围对变压器越限情况进行宏观分析，了解该区域的变压器负载情况，为电网的规划、系统调度和变压器运行检修等提供有效依据。

文献[4]以广东电网专用变压器为分析对象，提取最大负载率、平均负载率和负载率斜率等指标作为变压器经济运行的特征量，以历史数据为样本，聚类分析了广东电网专用变压器运行情况，并对分群结果进行了剖析，能够为宏观掌握广东电网专用变

压器的经济运行情况提供依据。但是，论文提取的特征指标体系更多是从变压器的经济性考虑，对于变压器的异常负载情况支撑不足。

本文建立了变压器越限特征指标体系，包括越限运行总时间、越限运行频次和平均越限运行时间；基于变压器负载历史数据，提取变压器越限特征样本，并通过 DBSCAN 聚类算法^[5-6]对变压器特征数据集进行分群，分析变压器各分群特征，从越限运行的维度分析变压器负载特性。

1 变压器越限特征指标

一方面，从变压器负载导则以及绝缘结构老化原理等可以得出，变压器越限运行幅度、持续时间等衡量越限严重程度的指标是分析变压器绝缘结构老化、寿命损失情况的重要依据，这也是反映变压器负载特性的直观特征量；另一方面，变压器越限发生的频次、平均每次越限时间等与地区负载水平、变压器运行策略等息息相关，能在较大程度上反映变压器运行的外部环境。因此，本文提取越限频次、越限总时间以及平均越限时间等作为变压器的越限特征指标。

1) 越限频次：统计时段内越限发生的次数，设为 n 。其中，变压器从负载率超过 80% 开始，至首次回落 80% 以下记为一次越限，两次负载率数据的时间差记为本次越限时间。

2) 越限总时间：统计时段内越限运行的总时间，单位为分钟，记为 T_{all} ，设统计时间段内变压器共计越限 n 次，任意一次越限事件 i 的持续时间为 t_i ，

则

$$T_{all} = \sum_{i=1}^n t_i \tag{1}$$

3) 越限平均时间：统计时段内越限运行的平均时间，即越限总时间与越限频次的比值，单位为分钟，设为 T_{avg} ，则

$$T_{avg} = \frac{T_{all}}{n} \tag{2}$$

2 越限特征分析方法

根据变压器负载历史数据，提取变压器的越限特征，并建立特征样本集，采用聚类分析算法对样本集进行分群，分析、识别各分群变压器越限情况，主要步骤包括特征样本提取、聚类分群和结果分析等。

2.1 建立样本集

本文选取了 32 台 500 kV 主变压器为分析对象。根据经验对错误数据进行预处理，剔除负载率大于 110%以及单次越限时间 t_i 大于 1 000 min 的数据。然后，计算、建立 32 台主变压器的负载越限特征样本集，如表 1 所示。

表 1 负载越限特征样本集

变压器编号	越限运行频次	越限运行总时间/min	平均越限运行时间/min
1	73	10 069.27	137.935 2
2	55	7 311.267	132.932 1
3	45	6 244.767	138.772 6
4	54	7 571.083	140.205 3
⋮	⋮	⋮	⋮
31	75	9 906.967	132.092 9
32	5	1 188.667	237.733 3

表 2 标准化后特征样本集

变压器编号	标准化后的越限运行频次	标准化后的平均越限运行时间	标准化后的越限运行总时间
1	0.382 040 02	0.897 640 825	0.492 660 186
2	0.106 281 81	0.794 349 612	0.184 076 103
3	-0.046 917 195	0.914 930 146	0.064 748 719
4	0.090 961 91	0.944 508 4	0.213 146 185
⋮	⋮	⋮	⋮
31	-0.675 033 118	-0.774 525 672	-0.608 474 787
32	-0.659 713 218	2.958 042 918	-0.500 962 66

由负载越限特征样本集可见，三个特征之间的数量级相差很大，为了消除数量级数据带来的影响，需要对负载越限特征样本集进行标准化处理，具体如下。

标准化后的越限运行总时间

$$Normtotal_t_i = \frac{total_t_i - \overline{total_t}}{\sqrt{Var_total_t}} \tag{3}$$

式中， $\overline{total_t} = \frac{1}{n} \sum_{i=1}^n total_t_i$ ；

$$Var_total_t = \frac{1}{n} \sum_{i=1}^n (total_t_i - \overline{total_t})^2。$$

标准化后的越限运行频次

$$Normnum_i = \frac{num_i - \overline{num}}{\sqrt{Var_num}} \tag{4}$$

式中， $\overline{num} = \frac{1}{n} \sum_{i=1}^n num_i$ ； $Var_num = \frac{1}{n} \sum_{i=1}^n (num_i - \overline{num})^2。$

标准化后的平均越限运行时间

$$Normavg_t_i = \frac{avg_t_i - \overline{avg_t}}{\sqrt{Var_avg_t}} \tag{5}$$

式中， $\overline{avg_t} = \frac{1}{n} \sum_{i=1}^n avg_t_i$ ； $Var_avg_t = \frac{1}{n} \sum_{i=1}^n (avg_t_i - \overline{avg_t})^2。$

经过标准化后的特征样本集见如表 2 所示。

2.2 聚类分析

聚类是一种常见的数据分析工具，其目的是将大量数据点的集合分成若干类，使每类中的数据之间最大程度地相似，而不同类中的数据最大程度地不同。通过聚类算法能够发现数据的不用群组，进而能够分析不同群组的内部特征。传统的基于层次聚类和划分聚类算法受噪声的影响大，聚类结果不稳定。而基于密度的聚类算法(如 DBSCAN 算法)，使用簇的基于密度的定义，因此它是相对抗噪声的，并且能够处理任意形状和大小的簇，且可以发现使用 K 均值不能发现的许多簇。

本文采用基于密度的聚类算法 DBSCAN 进行特征分群。为了更简洁描述 DBSCAN 聚类算法，先给出以下定义：

- 1) 一个给定对象的 Eps 半径内的区域称为该对象的 Eps 邻域。
- 2) 如果一个对象的 Eps 邻域内至少包含 MinPts 个对象，则称为核心对象。
- 3) 对于一个给定对象，如果它在某个核心对象的 Eps 邻域内，但自己不是核心对象，则称其为边界对象。
- 4) 给定对象集合，如果 q 在 p 的 Eps 邻域内，而 p 是核心对象，则称 p 到 q 是直接密度

可达。

- 5) 对于给定的 Eps 和 MinPts，如果存在一系列对象 $p_1,...,p_n,p_1=p,p_n=q$ ，其中 p_i 可以直接密度可达 p_{i+1} ，则称 p 到 q 是密度可达。

利用 DBSCAN 聚类算法，距离函数选择欧氏距离，选择合适的半径 Eps 和邻域内最小样本点数目 MinPts，对标准化后的特征样本集聚类分群，具体步骤如下：

- 1) 在特征样本集中随机选择一个不属于任何分群的核心对象变压器 p ，创建一个新的分群。
- 2) 根据该分群中的核心对象变压器，循环收集密度可达的核心对象变压器加入该分群，直到没有新的核心对象变压器加入为止。
- 3) 若不存在不属于任何聚类的核心对象变压器，则转入步骤 4)；否则返回步骤 1) 继续执行。
- 4) 对各个核心对象变压器收集与其密度相连的边界对象变压器，将边界对象变压器归类到相应核心对象变压器所属的聚类中。

下面对标准化后的特征样本集进行 DBSCAN 聚类，选择半径 Eps=0.7 和邻域内最小样本点数目 MinPts=2，其输出的聚类结果如表 3 和表 4 所示。

表 3 聚类结果（一）

分群类别	分群 0	分群 1	分群 2
样本个数	4	11	17
样本个数占比（%）	12.50	34.38	53.13

表 4 聚类结果（二）

变压器编号	越限运行频次	平均越限运行时间	越限运行总时间	分群号
1	73	137.935 2	1 0069.27	1
2	55	132.932 1	7 311.267	1
3	45	138.772 6	6 244.767	1
4	54	140.205 3	7 571.083	1
⋮	⋮	⋮	⋮	...
31	4	56.9416 7	2 27.7667	2
32	5	237.733 3	1 188.667	0

3 结果分析

根据 DBSCAN 聚类分群结果，对各个分群

分别针对负载越限特征指标，采用核密度估计模拟其概率密度函数，并画出相应的概率密度函数图像，获取负载越限特征指标值的区间范围，进

一步分析各个分群不同负载越限特征指标的变压器分布情况，评价不同分群的越限运行状态。

(1) 分群 0 特征分析

分群 1 的三个特征变量的频率密度如图 1 所示。

从图 1 可以看出，对于分群 0，变压器的越限运行频次分布不均，分布范围广，大致在

0~400 之间；平均越限运行时间较长，大致在 100~250 min；越限运行总时间较长，大致在 0~50 000 min。总体来看，分群 0 的越限严重程度较高，需重点关注并进一步分析越限原因等。

(2) 分群 1 特征分析

分群 2 的三个特征变量的频率密度如图 2 所示。

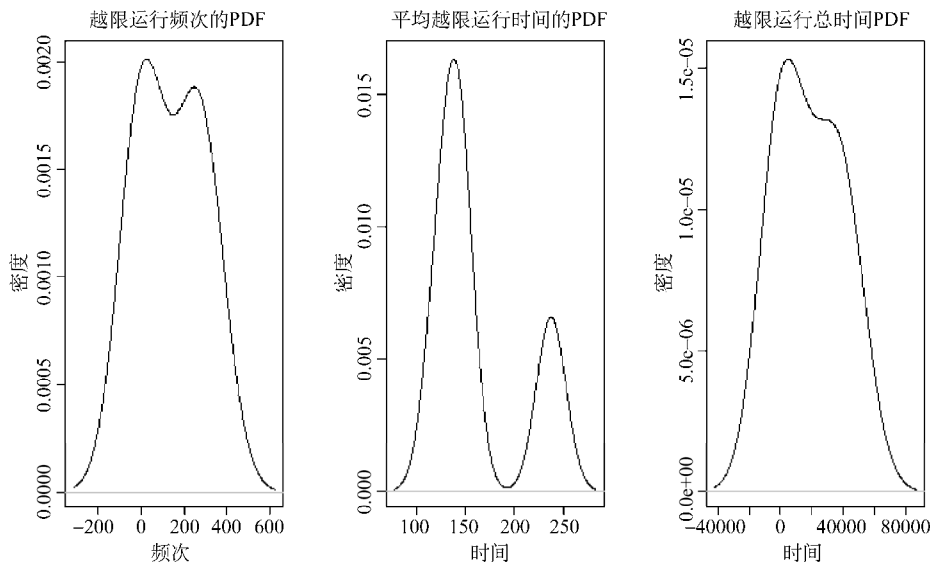


图 1 分群 0 特征指标的概率密度

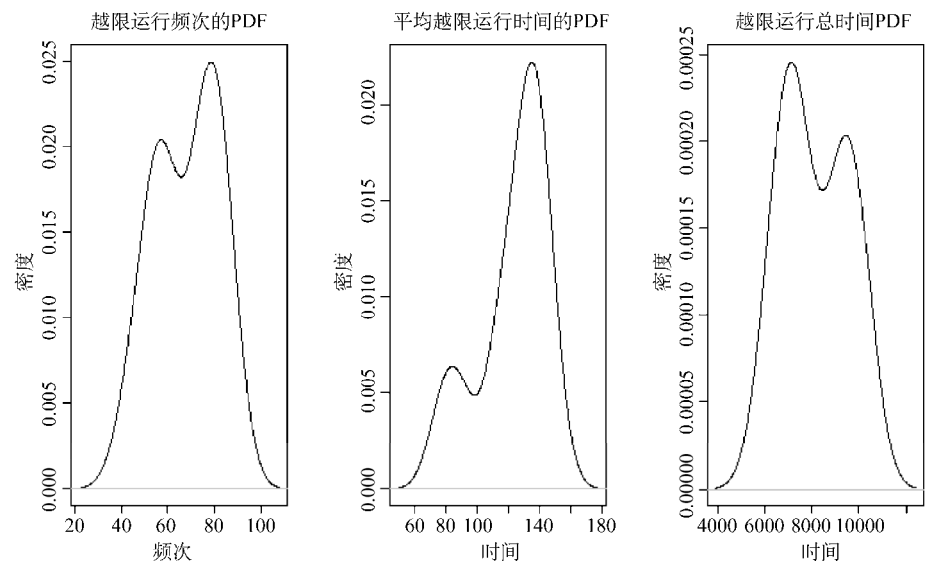


图 2 分群 1 特征指标的概率密度

从图 2 可以看出，对于分群 1，变压器的越限运行频次处于中等水平，大致在 20~100 之间；平均越限运行时间处于中等水平，大致在 70~150 min；越限运行总时间处于中等水平，大

致在 5 000~10 000 min。总体来看，分群 1 的越限严重程度处于中等水平。

(3) 分群 2 特征分析

三个特征变量的频率密度如图 3 所示。

从图 3 可以看出,对于分群 2,变压器的越限运行频次较低,大致在 0~40 之间;平均越限运行时间较短,大致在 0~110 min;越限运行总

时间较短,大致在 0~3 000 min。总体来看,分群 2 的越限严重程度较低。

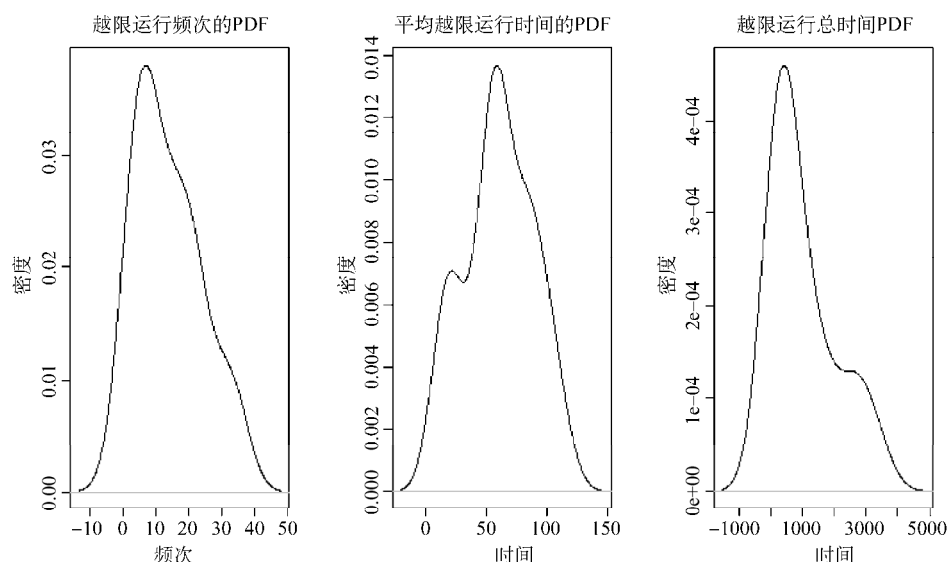


图 3 分群 2 特征指标的概率密度

上述变压器分群结果,对于电网规划、系统调度和变压器运行检修等具有直接指导价值。比如对于分群 0 所属变压器,在日常巡检、专业检修等工作中应重点关注,并可以进一步针对性地分析每台变压器的负载情况及运行方式等,适当考虑增加主变压器容量或改变运行方式,实现提高电网安全稳定水平以及变压器经济运行水平等。

4 结束语

本文提出了变压器越限特征分析方法,并建立特征指标体系,利用变压器的负载数据,提取变压器的越限特征指标样本,通过 DBSCAN 聚类算法对变压器特征数据集进行分群分析。通过对各个分群的具体分析,可以为安排变压器的运行方式、电网规划等提供参考价值。

参考文献

- [1] 李超. 变压器动态经济运行与降损分析[D]. 济南: 山东大学, 2011.
- [2] 周虹妤. 利用变压器中压侧无载档合理控制电压

- 越限时间[J]. 广西电力, 2011, 34 (1): 44-45.
- [3] 电力变压器 第 7 部分: 油浸式电力变压器负载导则[S]. 2008.
- [4] 刘涛, 杨劲锋, 等. 基于聚类的大量变压器负载状态分析[J]. 机电工程技术, 2013, 42 (9): 51-55.
- [5] 贺玲, 吴玲达, 蔡益朝. 数据挖掘中的聚类算法综述[J]. 计算机应用研究, 2007 (1): 10-13.
- [6] 冯少荣, 肖文俊. DBSCAN 聚类算法的研究与改进[J]. 中国矿业大学学报, 2008, 37 (1): 105-111.
- [7] 李莉平, 沈俊媛. 基于数据挖掘的 DBSCAN 算法及其应用[J]. 科技创业月刊, 2009 (8): 134-135.

作者简介

杜双育 (1985-), 男, 工学博士, 主要研究方向为电力设备在线监测。

范颖 (1980-), 女, 高级工程师, 工学硕士, 主要研究方向为电力设备状态监测及故障诊断。

(收稿日期: 2015-12-10) EA

(编辑: 胡颖)