



用. 文献[2]提出了一种基于支持向量机和决策函数的暂态稳定评估方法. 该方法以故障前潮流量为初始特征集, 结合暂态安全稳定量化评估和统计理论方法, 提取输入特征; 通过支持向量机训练暂态稳定评估模型, 得出评估模型的决策函数, 并依据支持向量的决策值确定阈值, 保证评估结果保守性. 文献[3]提出了一种结合上一时间断面评估结果和电网实时运行方式变化的安全稳定裕度快速估算方法.

国外在机器学习人工智能方法在电力系统中应用方面也进行了大量的研究. 文献[4]通过对几种机器学习回归方法的研究和比较, 建立了一种预测模型, 可以预测联合循环电厂每小时满负荷发电量. 文献[5]提出一种基于神经网络的用于中压/低压变电站负荷预测的机器学习模型. 文献[6]提出一种利用谱能量和机器学习作为人工神经网络, 利用  $k$  近邻和支持向量机对信号进行分类方法. 文献[7]提出了一种基于深度神经网络的特征提取方法.

本文融合了传统的因果分析方法和大数据思维<sup>[8]</sup>方法, 提出一种基于因果分析结果构建样本库, 基于机器学习进行稳定裕度快速估算的方法, 提升分析计算的速度, 为电网安全稳定快速决策提供依据, 为电力系统暂态稳定分析评估提供了新思路.

## 1 技术框架

技术框架如图1所示.

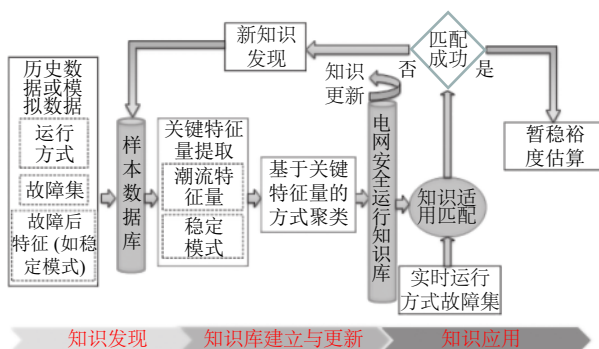


图1 基于机器学习的电网暂态稳定裕度估算技术框架图

首先从海量的历史量化分析或模拟仿真计算结果中, 按故障提取暂态稳定模式, 将所有故障的稳定模式取并集作为关键特征量. 这种电网暂态稳定关键特征量的提取方法, 从高维的数据中提取与电网暂态稳定性密切相关的关键特征量, 有效降低属性维数.

其次对各故障按特征量进行方式聚类, 生成电网安全运行知识库. 将海量历史运行方式聚类成为若干个运行方式簇, 同一运行方式簇内的数据具有比较接近的运行方式, 并且针对同一考核故障下具有相同的暂态稳定模式.

最后基于知识库对当前方式下多重故障进行自动匹配和稳定裕度快速估算. 比较当前实时运行方式与各聚类基准方式的关键设备投停状态是否一致, 如果一致则计算与该聚类基准方式的距离, 选取距离大于阈值且最小的方式簇, 从而得到当前实时运行方式所属的方式簇, 并分析当前实时运行方式与历史运行方式之间的差异, 采用基于历史分析结果的暂态稳定裕度快速估算方法, 快速估算出当前实时运行方式的暂态稳定裕度.

如果当前实时运行方式未匹配到任何历史方式簇, 则必须对当前电网实时运行方式进行详细的仿真分析, 将运行方式及仿真结果纳入到历史样本库中.

## 2 特征值提取

在机器学习的实际应用中, 特征数量往往较多, 其中可能存在不相关的特征, 特征之间也可能存在相互依赖或冗余, 特征数量越多, 容易导致以下问题:

- 1) 分析特征、训练模型所需的时间增加, 模型更复杂.
- 2) 容易引起“维度灾难”, 其推广能力会下降.
- 3) 容易导致机器学习中经常出现的特征稀疏的问题, 导致模型效果下降.
- 4) 对于模型来说, 可能会导致不适定的情况, 即使解出的参数会因为样本的微小变化而出现大的波动.

特征选择是指从已有的  $M$  个特征 (Feature) 中选择  $N$  个特征使得系统的特定指标最优化, 是从原始特征中选择出一些最有效特征量以降低数据集维度的过程, 是提高学习算法性能的一个重要手段, 也是模式识别中关键的数据预处理步骤<sup>[9-12]</sup>.

从在线安全分析应用长期运行积累的海量历史数据或构造的海量历史数据中, 获取的电网运行方式数据信息包括电网的运行状态量, 主要包括: 母线电压幅值、母线电压相角、发电机有功无功、负荷有功无功、交流线路有功无功、变压器有功无功、直流线路有功无功、容抗器投入无功容量. 基于 EEAC 的安全稳定量化分析方法能够给出各故障下的暂态功角稳定

裕度、暂态电压安全裕度、暂态频率安全裕度、发电机分群模式,各发电机对暂态功角稳定的参与因子,各母线对暂态电压稳定的参与因子,母线或发电机对暂态频率稳定的参与因子等信息.可以根据参与因子对电网的运行状态量进行筛选,提取与电网暂态稳定性密切相关的关键特征量,有效降低特征维数.选择参与因子大于设定门槛值的相关元件状态变量作为关键特征量,具体包括:

(1) 在加速群机组中,选择机组参与因子大于设定门槛值 $\lambda_{set-A}$ 的机组状态变量作为关键特征量;

(2) 在减速群机组中,选择参与因子绝对值大于设定门槛值 $\lambda_{set-S}$ 的机组状态变量作为关键特征量;

(3) 在加速群节点中,选择负荷参与因子大于设定门槛值 $\lambda_{set-NA}$ 的负荷节点状态变量作为关键特征量;

(4) 在减速群节点中,选择负荷参与因子绝对值大于设定门槛值 $\lambda_{set-NS}$ 的负荷节点状态变量作为关键特征量;

$\lambda_{set-A}$ 、 $\lambda_{set-S}$ 、 $\lambda_{set-NA}$ 、 $\lambda_{set-NS}$ 需要根据实际工程情况确定,一般 $\lambda_{set-A}$ 、 $\lambda_{set-S}$ 取值0.15, $\lambda_{set-NA}$ 、 $\lambda_{set-NS}$ 取值0.3.

### 3 方式聚类

聚类分析指将物理或抽象对象的集合分组为由类似的对象组成的多个类的分析过程.聚类就是按照某个特定标准(如距离准则,即数据点之间的距离)把一个数据集分割成不同的类或簇,使得同一个簇内的数据对象的相似性尽可能大,同时不在同一个簇中的数据对象的差异性也尽可能地大.可以具体地理解为,聚类后同一类的数据尽可能聚集到一起,不同类数据尽量分离.

实际电网运行方式具有明显的规律性和重复性,对历史数据进行分析 and 处理,按照关键特征量将海量历史运行方式聚类成为若干个运行方式簇,同一运行方式簇内的数据具有比较接近的运行方式,并且同一考核故障下具有相同的暂态稳定模式.通过对电网运行方式的聚类,可以有效提取电网中的典型运行方式,确保电网安全分析的完备性,有利于发现孤立场景,提高电网运行分析精度和电网运行管理决策的科学性,实现运行方式筛选的智能化.

考虑电网暂态稳定模式的电网历史运行方式聚类方法(图2),通过对历史数据进行分析 and 处理,按照电

网的暂态稳定模式,将海量历史运行方式聚类成为若干个运行方式簇,同一运行方式簇内的数据具有比较接近的运行方式,并且针对同一考核故障下具有相同的暂态稳定模式.每一簇内包含多个运行方式接近的历史运行方式,也可能只包括一个历史运行方式,具体包括以下步骤:

(1) 获取全部电网历史运行方式数据、故障信息,以及各故障下对应的暂态稳定结果信息,包括加速群机组、减速群机组、参与因子等稳定模式信息,以及暂态稳定裕度;

(2) 针对考核故障 $F_m$ ,对包含该故障下的全部历史运行方式进行聚类;设定每个历史方式下的详细仿真得到的暂态功角裕度为 $\eta_i$ ,根据加减速机组各自参与因子,筛选出加速机组集合为 $G_{si}$ ,减速机组集合为 $G_{ai}$ ,特征线路投运集合 $L_i$ ;

(3) 初始时选择一个方式 $K$ 作为基准方式,加入基准方式集合;

(4) 针对所有方式 $i$ ,判断该考核故障的加速机组集合 $G_{si}$ 、减速机组集合 $G_{ai}$ 和特征线路投运集合 $L_i$ 是否和基准方式集合的基准方式的加速机组集合 $G_{sk}$ 、减速机组集合 $G_{ak}$ 以及特征线路投运集合 $L_k$ 一致,即:

- ① 加速机组个数和机组名称和基准方式一致;
- ② 减速机组个数和机组名称和基准方式一致;
- ③ 特征线路个数、投运方式和名称和基准方式一致.

若上述条件均满足,则转(4),如果与基准方式集合所有的基准方式都不一致,则将该方式作为一新的基准方式加入到基准方式集合,处理下一个方式;

(5) 根据待分类的方式 $i$ 与基准方式 $K$ 之间的方式差异,以及基准方式 $K$ 的暂态稳定裕度 $\eta_k$ ,利用方式差异和历史分析结果的暂态稳定裕度快速估算方法,计算待分类方式 $i$ 的裕度 $\eta_{i-e}$ ;

(6) 设定暂态功角稳定裕度误差上限 $\Delta\eta$ (裕度误差根据实际需求设置,如10.0),若满足以下条件时:

- ①  $|\eta_{i-e} - \eta_i| \leq \Delta\eta$ ;
- ②  $\eta_{i-e} \times \eta_i > 0$ (即 $\eta_{i-e}$ 、 $\eta_i$ 均为正或均为负).

则认为待分类的方式 $i$ 与基准方式 $K$ 可聚成同一簇,如果基准方式集合无法匹配到同类基准方式,则将该方式作为一新的基准方式加入到基准方式集合;处理下一个方式,如果所有方式聚类完成,转步骤(7);

(7) 将同簇的数据各方式数据之间根据加减速机组功率和参与因子乘积的值序列的差值,找到该簇的

中心点,将离中心点最近的方式作为该簇的基准方式;  
(8) 如果基准方式不再变化或者达到最大迭代聚

类次数结束,否则,按新的基准方式形成基准方式集合,重复步骤(4)至步骤(7)。

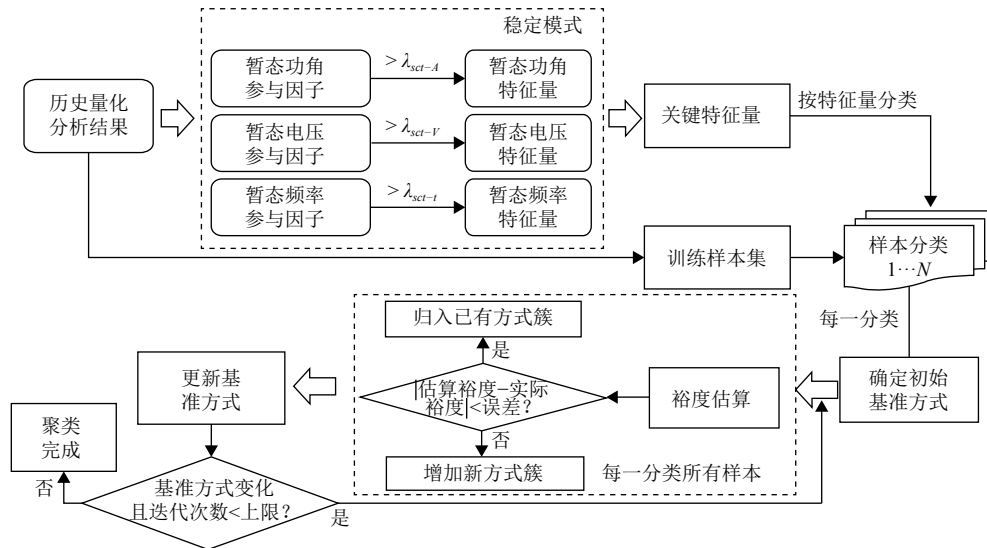


图2 方式聚类流程图

#### 4 方式分类和裕度估算

基于前述的关键特征量的选择方法,通过关键特征量找到与当前电网实时运行方式接近的历史运行方式簇,进一步根据当前电网实时运行方式与历史运行方式之间的差异,结合历史运行方式的暂态稳定裕度情况,快速估算当前电网实时运行方式的暂态稳定裕度,具体步骤如下:

1) 根据考核的故障,比较当前实时运行方式与各聚类基准方式的关键设备投停状态是否一致,如果一致则计算与该聚类基准方式的加权距离,选取距离大于阈值且最小的方式簇,从而得到当前实时运行方式所属的方式簇.如果当前实时运行方式未匹配到任何历史方式簇,则必须对当前电网实时运行方式进行详细的仿真分析.

2) 从匹配到的历史运行方式簇,选取其中历史运行方式及其暂态稳定分析结果,根据在线方式与历史运行方式之间的关键特征量的差异,采用基于历史分析结果的暂态稳定裕度快速估算方法,估算当前电网运行方式的暂态稳定裕度,估算方法参见文献[3].

对于特征值向量  $X$ ,由于每个分量对于稳定性的影响是不同的,因此,在计算当前方式与基准方式的距离时需要考虑每个分量的权重.实际分配权重有两种方式,一是以基准方式各状态量的参与因子作为权重,

二是以该聚类所有方式状态量参与因子的均值作为权重.对应的距离公式分别为:

$$D(X, Y) = \sqrt{\gamma_1 x_1 - y_1^2 + \gamma_2 x_2 - y_2^2 + \dots + \gamma_n x_n - y_n^2} \quad (1)$$

$$D(X, Y) = \sqrt{\bar{\gamma}_1 x_1 - y_1^2 + \bar{\gamma}_2 x_2 - y_2^2 + \dots + \bar{\gamma}_n x_n - y_n^2} \quad (2)$$

式中,  $X$  为当前方式特征值向量,  $Y$  为基准方式特征值向量,  $\gamma_n$  为基准方式各状态量  $y_n$  对应的参与因子,  $\bar{\gamma}_n$  为该聚类所有方式状态量  $y_n$  对应的参与因子均值.

#### 5 工程实例

以某省级电网为例,选择共 12 256 个运行方式及其计算结果构建训练样本,其中安全样本(裕度大于 10)10 818 个,预警样本(裕度大于 0 且小于 10)68 个,不安全样本(裕度小于 0)1370 个.选择共 851 个运行方式作为测试集.针对指定的测试故障进行裕度快速评估,训练样本分布图如图 3.



图3 训练样本分布图

根据前述的基于量化分析的暂态稳定特征量选择方法,提取包括116台机组有功值、7条线路的有功值共123个特征量作为暂态功角稳定的关键特征量。

根据前述的电网历史运行方式聚类方法,对测试故障下的共12256个训练样本集历史运行方式进行历史方式聚类,其中,聚类的依据为:

(1) 根据量化分析结果,测试故障下对应的加速群机组投运状态一致;

(2) 该故障对应的关键特征量中,选取的重要联络线投运状态一致;

(3) 聚类的暂态功角稳定裕度阈值设为 $\pm 5.0\%$ ,即各簇内的暂态功角稳定裕度最大差值不超过10%。

经前述考虑电网暂态稳定模式的电网历史运行方式聚类方法,12256个训练样本集共聚类成462个簇,其中部分簇内包含的历史样本较少。方式聚类分布图如图4。

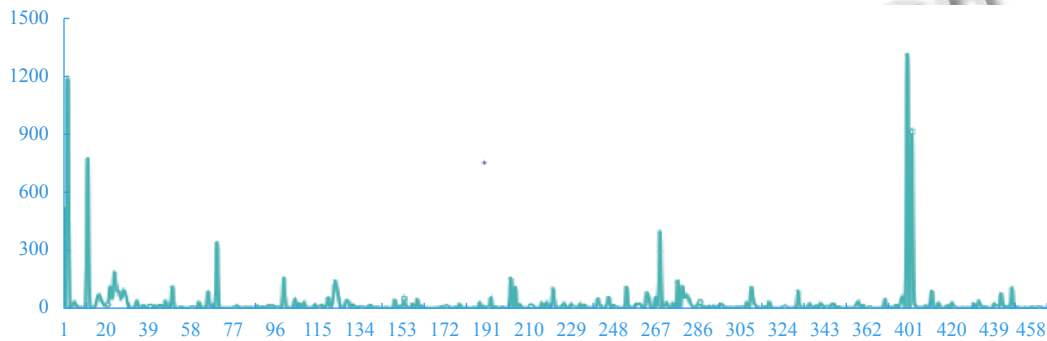


图4 方式聚类分布图

分别比较测试集中各方式与各聚类簇中历史方式的加速群机组投运状态、重要联络线投运状态,在加速群机组投运状态、重要联络线投运状态均一致的若干个簇中,计算测试集方式与簇内运行方式之间的方式距离,作为测试集方式的目标方式。进一步根据测试集方式与目标方式之间的运行方式差异,利用基于方式差异和历史分析结果的暂态稳定裕度快速估算方法,得出测试集方式的裕度及其稳定性结论。

将在线快速评估估算结果与实际仿真结果进行比较,测试集共851个样本。

按照式(1)计算距离进行方式匹配,结果如图5和图6。



图5 快速评估估算结果与实际仿真结果裕度比较结果1

共有34个测试样本出现错判,判断准确率为96%(817/851),误判率为4%(34/851),其中:



图6 快速评估估算结果与实际仿真结果裕度误差统计1

(1) 实际时域仿真分析结果为暂态功角稳定,快速评估估算结果判为失稳的共28个样本,误警率3.3%(28/851);

(2) 实际时域仿真分析结果为暂态功角失稳,快速评估估算结果判为稳定的共6个样本,漏警率0.70%(6/851)。

统计在线快速评估估算结果与实际仿真结果之间暂态功角稳定裕度结果之间的差异,两者之间的裕度平均绝对误差为16.94%,相对误差为1%。

按照式(2)计算距离进行方式匹配,结果如图7和图8。

共有6个测试样本出现错判,判断准确率为99.3%(845/851),误判率为0.7%(6/851),其中:

(1) 实际时域仿真分析结果为暂态功角稳定,快速评估估算结果判为失稳的共0个样本,误警率0%

(0/851);

(2) 实际时域仿真分析结果为暂态功角失稳, 快速评估估算结果判为稳定的共 6 个样本, 漏警率 0.70% (6/851).

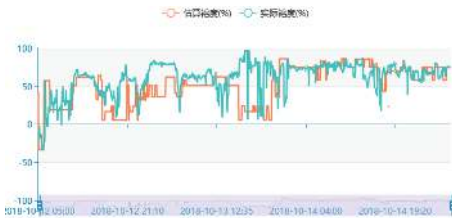


图7 快速评估估算结果与实际仿真结果裕度比较结果 2



图8 快速评估估算结果与实际仿真结果裕度误差统计 2

统计在线快速评估估算结果与实际仿真结果之间暂态功角稳定裕度结果之间的差异, 两者之间的裕度平均绝对误差为 18.01%, 相对误差为 0.66%.

从误差率的角度来看, 显然用式 (2) 比式 (1) 效果更好.

## 6 结论与展望

本文提出了一种安全稳定裕度快速估算方法. 首先基于海量历史评估结果和安全稳定机理提取与电网暂态稳定性密切相关的关键特征量, 其次然后对各故障按特征量进行方式聚类, 生成电网安全运行知识库, 最后基于知识库对当前方式下故障进行自动匹配和裕度快速估算. 通过具体工程实例说明上述方法的工程应用步骤和实际效果. 由于方式聚类时用到了评估结果中的裕度信息, 而新方式匹配时缺少这一信息, 工程上用特征量之间的加权距离进行匹配, 这可能是结果误差的主要来源, 不同距离计算方法会影响估算结果

的准确性, 后续还要进一步分析和研究.

## 参考文献

- 孙宏斌, 黄天恩, 郭庆来, 等. 基于仿真大数据的电网智能型超前安全预警技术. 南方电网技术, 2016, 10(3): 42-46.
- 徐泰山, 鲍颜红, 杨莹, 等. N-2 组合故障集的暂态功角稳定在线快速评估. 电力系统保护与控制, 2015, 43(7): 122-126. [doi: 10.7667/j.issn.1674-3415.2015.07.019]
- 徐泰山, 段荣华, 鲍颜红, 等. 基于预想故障集自动筛选的在线暂态安全稳定评估方法: 中国, CN201710247481.X. [2017-08-25].
- Tüfekci P. Prediction of full load electrical power output of a base load operated combined cycle power plant using machine learning methods. International Journal of Electrical Power & Energy Systems, 2014, 60: 126-140.
- Ding N, Benoit C, Foggia G, et al. Neural network-based model design for short-term load forecast in distribution systems. IEEE Transactions on Power Systems, 2016, 31(1): 72-81. [doi: 10.1109/TPWRS.2015.2390132]
- Morales J A, Orduña E, Rehtanz C. Classification of lightning stroke on transmission line using multi-resolution analysis and machine learning. International Journal of Electrical Power & Energy Systems, 2014, 58: 19-31.
- Bosse S, Maniry D, Müller KR, et al. Deep neural networks for no-reference and full-reference image quality assessment. IEEE Transactions on Image Processing, 2018, 27(1): 206-219. [doi: 10.1109/TIP.2017.2760518]
- 李国杰. 大数据研究的科学价值. 中国计算机学会通讯, 2012, 8(9): 8-15.
- 顾雪平, 曹绍杰, 张文勤. 神经网络和短时仿真结合的暂态安全评估事故筛选方法. 电力系统自动化, 1999, 23(8): 16-19. [doi: 10.3321/j.issn:1000-1026.1999.08.004]
- 黄天恩, 孙宏斌, 郭庆来, 等. 基于电网运行大数据的在线分布式安全特征选择. 电力系统自动化, 2016, 40(4): 32-40. [doi: 10.7500/AEPS20150424003]
- 国家电网公司, 国网湖南省电力公司, 中国电力科学研究院. 基于历史数据的电力系统稳定性快速判断方法: 中国, CN201511030190.2[P]. [2016-03-16].
- 蒋维勇, 孙宏斌, 张伯明, 等. 电力系统精细规则的研究. 中国电机工程学报, 2009, 29(4): 1-7. [doi: 10.3321/j.issn:0258-8013.2009.04.001]