

基于 GBDT 的串联结构微电网 与电网交互稳定性评估

张津京¹, 王兴贵^{1†}, 宋汶秦², 杨维满¹

(1. 兰州理工大学 电气工程与信息工程学院, 甘肃 兰州 730050;
2. 国网甘肃省电力公司 经济技术研究院, 甘肃 兰州, 730050)

摘要:针对微源逆变器串联微电网中,微源投切对微电网与电网交互稳定性影响机理复杂,传统方法建模困难且无法及时决策的问题,提出一种基于梯度提升决策树(gradient boosting decision tree, GBDT)模型的交互稳定性评估方法.首先提出微源投切特征函数的构造方法,所提方法能够有效提取采用载波移相 SPWM 调制的串联结构系统中特有的投切微源位置特征,同时具有降低特征维数、提高模型效率、通过简单拟合使模型具备迁移能力的优势;其次将交互稳定性评估视为多分类任务,且对其中较不稳定样本进行谐振频段预测,从而可针对不同投切工况采取不同应对策略.在验证集和新建测试集上的测试结果表明特征提取方法正确、有效,可为类似串联结构系统在电能质量分析、优化运行等方面提供方法参考;所建模型泛化能力强且具有实用性,可为并网运行系统稳定性在线分析及决策提供新思路.

关键词:微源投切;微源逆变器串联;交互稳定性;特征提取;机器学习

中图分类号:TM61

文献标志码:A

Assessment of Interactive Stability between Grid and Series Structure Micro-grid Based on GBDT

ZHANG Jinjing¹, WANG Xinggui^{1†}, SONG Wenqin², YANG Weiman¹

(1. School of Electrical Engineering and Information Engineering, Lanzhou University of Technology, Lanzhou 730050, China;
2. State Grid Gansu Electric Power Corporation Economical & Technical Research Institute, Lanzhou 730050, China)

Abstract: Due to the complex mechanism of the effect of micro-source switching on interactive stability between the grid and the micro-grid with series micro-source inverters, traditional methods were difficult to model and can't make decision in time. In this study, therefore, an assessment method of interactive stability based on gradient boosting decision tree (GBDT) was proposed. Firstly, the feature functions of the micro-source switching were constructed to extract the position feature which existed typically in series structure system modulated by carrier phase shifted SPWM. The proposed functions can reduce the feature dimension, thus improving the efficiency of assessment model and making the model migration capable by feature fitting. Secondly, the assessment of interactive stability was treated as

* 收稿日期:2020-01-16

基金项目:国家自然科学基金资助项目(51867016), National Natural Science Foundation of China(51867016); 国家电网公司科学技术项目(52272817000L), Science and Technology Project of State Grid Corporation of China (52272817000L); 甘肃省高等学校创新能力提升项目(2019A-024), Innovation Ability Improvement Project of Colleges and Universities in Gansu Province(2019A-024)

作者简介:张津京(1983—),女,甘肃金昌人,兰州理工大学博士研究生

† 通信联系人, E-mail: wangxg689@126.com

a multi-classification task and the resonant frequency bands of instable samples were predicted. Thus, the different strategy can be taken according to the classification and prediction. Finally, the tests on validation data sets and new test sets verified the correctness and effectiveness of the feature construction functions which can provide references for the similar series structure system in power quality analysis, optimal operation and other aspects. The results also verified the high generalization performance and practicality of the proposed assessment model, which can provide a new tool for on-line stability analysis and decision-making of grid-connected operation systems.

Key words: micro-source switching; series micro-source inverters; interactive stability; feature extraction; machine learning

微电网是分布式电源 (distributed generation, DG) 接入电网的一种有效手段. 相比于 DG 和负荷在馈线上横向分布的并联结构微电网, 本文讨论的微源逆变器串联微电网 (micro-grid with series micro source inverters, SMSI-MG) 输出的电压和电流具有更好的正弦度, 同一输出电压等级下可有效降低各微源直流侧电压^[1-3]. 类似的逆变器串联结构因其高输出电压、低 dv/dt 、高效率等优势而在光伏并网系统中获得广泛应用^[4-5]. SMSI-MG 系统则在此基础上将不同类型的 DG 进行整合, 可进一步提高可再生能源利用率.

目前, 微电网中的大部分电源均为使用并网变流器作为接口的 DG, 在并网运行时由于变流器的高频开关特性而产生大量谐波, 污染电网. 进一步地, 变流器与变流器之间、变流器与电网之间的谐波交互作用也可能引起谐振, 导致网侧电压/电流谐波幅值持续放大、严重畸变, 进而造成保护装置误动、变流器控制环失稳等, 严重影响系统的安全稳定运行^[6-7]. 这种变流器接入电网后的谐波不稳定现象早在 2014 年就引起了丹麦 Aalborg 大学 F. Blaabjerg 课题组的关注^[8-9]. 之后, 国内外学者对谐波振荡引起的交互系统稳定性问题进行了大量研究^[10-11], 但主要针对弱电网环境下的单台逆变器或多台逆变器并联的系统, 对 SMSI-MG 这种微源逆变器 (micro-source inverter, MSI) 串联的系统研究较少. SMSI-MG 中 MSI 与 MSI 之间、串联的 MSI 与电网之间同样存在谐波交互作用. 特别地, 当微源因故障或出力受限进行投切操作后, SMSI-MG 与电网之间的交互作用更加明显, 也更加复杂. 由于谐波振荡引起的交互系统不稳定具有一定的危害性, 因此对其提前进行估计和防范具有重要意义.

从研究方法来说, 目前对并网运行系统稳定性

进行分析时主要采用状态空间法^[12-13]和阻抗分析法^[14-15]. 然而时频域的分析方法存在模型复杂、计算速度慢、实时性差等问题, 无法在微源投切后短时间内进行分析决策. 近年来, 基于数据的机器学习在电力系统暂稳评估研究中取得了重大进展^[16-17]. 机器学习方法将稳定性评估看作一个分类问题, 通过离线自动学习建立物理量与稳定后果之间的映射关系. 集成学习方法则将几种机器学习技术组合成一个预测模型, 可减小方差, 改进预测效果^[18]. 其中, 基于树模型的集成学习方法梯度提升决策树 GBDT 更是在 Kaggle、KDD (knowledge discovery in database) 等一系列数据挖掘竞赛中受到广泛关注. 考虑到微源投切对 SMSI-MG 和电网交互稳定性影响机理不明, 传统方法无法及时决策等问题, 本文借鉴现有电力系统暂稳评估思想, 提出一种基于 GBDT 的交互稳定性影响评估方法. 在建立机器学习模型时, 特征提取是关键. 因此首先通过分析微源投切与系统输出电压之间的关系, 提出特征提取和构造方法. 该方法同时考虑到了系统拓扑结构发生变化时所建模型的适应能力, 通过简单地拟合即可使模型适应于新网络拓扑, 具备良好的泛化性能. 最后对所建模型进行了测试.

1 微源投切特征构建及稳定性分级

图 1 为并网运行模式下 SMSI-MG 系统拓扑. 风力、光伏等微源直流侧配备储能装置, 通过对其有效控制可抑制功率波动从而保证各 MSI 直流侧电压相等. H 桥 MSI 输出端串联, 系统输出的多电平电压经简单滤波后即可具有较好的正弦度. 正常运行时, 单相微源数目为 n . 三相系统由单相系统组合而成. 为描述方便, 将单相包含 n 个微源的三相 SMSI-MG 系

统简记为 n -SMG. 该系统在公共耦合点 (point of common coupling, PCC) 处接入电网. 采集 PCC 处电压电流经控制器和载波移相 SPWM 调制 (carrier phase shifted SPWM, CPS-SPWM) 后实现对各 MSI 的控制, 从而使系统输出满足电网调度和负荷需求. 微源投切通过开关 SC 实现.

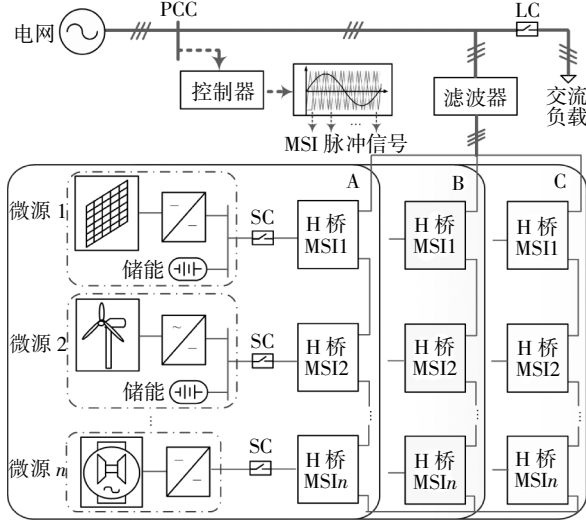


图1 并网运行模式 SMSI-MG 系统拓扑

Fig.1 System equivalent topology of grid-connected SMSI-MG

1.1 输入特征构建

利用机器学习训练交互稳定性评估模型, 就是要建立输入特征与交互稳定状态的映射关系, 因此输入特征的选取和构建是首要的步骤. 在对微源投切引起的交互稳定性影响进行评估时, 若将三相所有微源的投切状态作为特征属性, 则对于一个 n -SMG 系统, 仅该部分特征维数就为 $3n$. 遍历各种投切工况将产生“组合爆炸”问题, 而在设置训练样本时则易出现大量相似或是遗漏工况, 导致训练模型欠拟合、泛化性能差. 为避免上述问题, 本文对投切工况 $S_l (l=1, 2, \dots, L)$ 构造如下特征函数:

1) 三相微源平均投切率

投切微源的数目直接影响运行微源的数目, 进而影响系统输出电压中基波和谐波的含量. 为使模型具备迁移能力, 定义投切工况 S_l 对应的 A 相切除微源比例为 δ_{ac} , 恢复投入的微源比例为 δ_{ap} , 则 $\delta_{ap} \leq \delta_{ac}$. 令 $\delta_a = \delta_{ac} - \delta_{ap}$ 为 A 相微源投切率, 其余两相同理. 由于 A、B、C 三相微源投切对交互系统稳定性的影响是等价的, 因此为了弱化具体相特征对模型泛化能力的影响, 定义工况 S_l 对应的三相微源平均投切率为

$$\alpha_l = (\delta_a + \delta_b + \delta_c) / 3 \quad (1)$$

2) 三相微源投切不平衡度

定义工况 S_l 对应的三相微源投切不平衡度为

$$\beta_l = \frac{\sqrt{(\delta_a - \delta_b)^2 + (\delta_b - \delta_c)^2 + (\delta_c - \delta_a)^2}}{2(\delta_a + \delta_b + \delta_c)} \quad (2)$$

将特征 α_l 和 β_l 结合既可获得 S_l 对应的三相投切微源数目信息, 同时也使特征维数由 3 个 $\delta_j (j$ 代表 A、B、C 三相, 下文中含义相同) 降为 2 个.

3) 三相微源投切平均离散度

单相 SMSI-MG 采用 CPS-SPWM 调制方式. 若相邻载波水平移相 $2\pi/N$, N 个两电平经过水平移相叠加后可得输出电压谐波表达式为^[19]

$$u_o = \frac{NM}{2} \cos(\omega_0 t) + \sum_{m=1}^{\infty} \sum_{n=\pm 1}^{\infty} \frac{2}{m\pi} J_n\left(m \frac{\pi}{2} M\right) \sin\left[(m+n) \frac{\pi}{2}\right] \cdot \sum_{i=1}^N \cos\left[(m\omega_c + n\omega_0)t + m(i-1) \frac{2\pi}{N}\right] \quad (3)$$

式中: M 为调制度; $J_n(\cdot)$ 为 n 阶 Bessel 函数; ω_0 为基波角频率, $n\omega_0$ 表示基波倍数的谐波; ω_c 为载波角频率, $m\omega_c$ 表示载波倍数的谐波.

将图 1 中 A 相微源按其 MSI 串联顺序依次标记为 MS1、MS2、 $\dots$ 、MS n , 若微源 MS i 被切除, 则第 i 个两电平输出为 0, 其余两相同理. 取 $N=10, m=1, n=2, t=0$, 假设仅 A 相微源进行表 1 中投切操作, 而 $\delta_b = \delta_c = 0$, 分别计算各 S_l 下式 (3) 中求和项 $\sum_{i=1}^N \cos\left[(m\omega_c + n\omega_0)t + m(i-1) \frac{2\pi}{N}\right]$ 的结果及方差如表 1 所示 (为分析方便, 对求和结果取绝对值).

表 1 不同投切工况下求和项计算结果

Tab.1 Calculation results of summation terms under different switching conditions

序号	S_l	切除微源	δ_a	$ \Sigma $	方差
①	S_1	MS1, MS2	0.2	1.809	0.46
	S_2	MS1, MS3	0.2	1.309	
	S_3	MS1, MS4	0.2	0.691	
	S_4	MS1, MS5	0.2	0.191	
	S_5	MS1, MS6	0.2	0	
②	S_1	MS2, MS1	0.2	1.809	0.44
	S_2	MS2, MS3	0.2	1.118	
	S_3	MS2, MS4	0.2	0.5	
	S_4	MS2, MS5	0.2	0	
	S_5	MS2, MS6	0.2	0.191	
③	S_1	MS1, MS2, MS3	0.3	2.118	0.91
	S_2	MS1, MS2, MS10	0.3	2.618	
	S_3	MS1, MS5, MS10	0.3	1	
	S_4	MS1, MS3, MS5	0.3	0.5	
	S_5	MS1, MS4, MS7	0.3	0.118	

表 1 中,数据①为切除 2 个微源的情况,其中一个为 MS1,另一个则相对 MS1 的位置不同.由式(1)、(2)可知 5 种 S_i 下 α_i 和 β_i 相同.但 5 组求和结果的方差较大,由于此时式(3)中其他变量相同,因此可知被切除微源之间的相对位置不同时, u_0 差异较大.数据②、③进一步说明了这一点.由此可以推断对于 n -SMG 系统,当 α_i 和 β_i 相同时,若投切微源之间的相对位置不同,则系统输出电压电流中的谐波含量差异较大,其对交互稳定性的影响程度必然不同.这是采用 CPS-SPWM 调制的串联结构系统中特有的现象.因此投切微源的位置应当作为一个与投切数目同样重要的输入特征参与交互稳定性的评估.

为提取工况 S_i 对应的 j 相投切微源之间的相对位置特征,首先构造投切数组 AR_{ij} .如图 2 所示,微源仍按其 MSI 串联顺序标记为 MS1、MS2、 $\cdots$ 、MS n .切除微源在图中用虚线框表示,如 A 相的 MS1、MS2 和 MS3.对应于 AR_{ij} 中,将其序号 i 列出,而未切除的微源则标记为 0,故此时 $AR_{ia} = [1, 2, 3, 0, 0, \cdots, 0]$,其余与之同理.在获得 AR_{ij} 后,构造对应的离散度函数用于表征投切微源之间的相对位置.该离散度函数应当满足:1) 投切微源之间相对位置变化规律相同则离散度相同;2) δ_i 相同的情况下,切除连续序号的微源越多则离散度越小.据此构造如下所示工况 S_i 对应的 j 相投切微源离散度函数为

$$f_j(AR_{ij}) = \begin{cases} 0 & n_{jc} = 0 \\ \sum_{i \in R_{ij}} (i - \mu_j)^2 / n_{jc}, & n_{js} = 0 \\ \sum_{i \in R_{ij}} (i - \mu_j)^2 / (2^{n_{js}} \cdot n_{jc}), & n_{js} \neq 0 \end{cases} \quad (4)$$

式中: R_{ij} 为工况 S_i 对应的 j 相投切微源集合; n_{jc} 为此投切微源的总数,定义 $\mu_j = (\sum_{i \in R_{jc}} i) / n_{jc}$ 为对应的投切微源中心,如图中星号所示; n_{js} 为投切微源序号连续的次数,即 $(AR_{ij}[i+1] - AR_{ij}[i] = 1)$ 的次数.

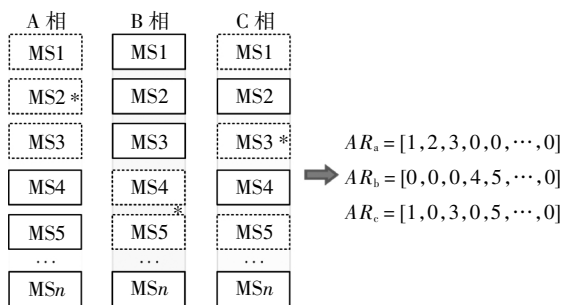


图 2 AR_{ij} 构造示意图

Fig.2 Diagram of constructing AR_{ij}

与特征 α_i 的考虑一样,定义三相平均离散度表

征工况 S_i 对应的投切微源相对位置,即

$$\gamma_i = (\sum_{j=a,b,c} f_j(AR_{ij})) / 3 \quad (5)$$

4) 三相微源投切中心欧氏距离

同样取 $N = 10, m = 1, n = 2, t = 0$,对 A 相微源进行投切操作,而 $\delta_b = \delta_c = 0$.计算若干投切微源之间相对位置相同时求和项 $\sum_{i=1}^N \cos[(m\omega_c + n\omega_0)t + m(i-1)\frac{2\pi}{N}]$ 的结果如表 2 所示.

表 2 $f_j(AR_{ij})$ 相同时求和项计算结果
Tab.2 Calculation results of summation terms when $f_j(AR_{ij})$ is the same

序号	S_i	切除微源	$f_a(AR_{ia})$	$ \Sigma $	方差
①	S_1	MS1, MS2	0.125	1.809	0.44
	S_2	MS3, MS4	0.125	0	
	S_3	MS5, MS6	0.125	1.809	
	S_4	MS7, MS8	0.125	1.118	
	S_5	MS9, MS10	0.125	1.118	
②	S_1	MS2, MS5	2.25	0	0.2
	S_2	MS3, MS6	2.25	0.691	
	S_3	MS4, MS7	2.25	1.118	
	S_4	MS5, MS8	2.25	1.118	
	S_5	MS6, MS9	2.25	0.691	
③	S_1	MS1, MS2, MS3	0.167	2.118	0.56
	S_2	MS3, MS4, MS5	0.167	0.809	
	S_3	MS5, MS6, MS7	0.167	2.618	
	S_4	MS6, MS7, MS8	0.167	2.118	
	S_5	MS8, MS9, MS10	0.167	0.809	

表 2 中,数据①为切除连续 2 个微源的情况,5 种 S_i 下被切除微源之间相对位置相同,因而由式(4)计算得到的离散度相同.数据②和③同理.但三组数据求和结果方差均较大,这说明即便投切微源数目相等、相对位置相同, u_0 还会因为投切微源的绝对位置不同而产生较大差异,因此需要进一步补充其他特征对其进行区分.观察表 2 中各 S_i 发现, α_i 、 β_i 和 γ_i 相同的情况下,投切微源中心 μ 是不同的,故利用离散度结合中心点的方式即可较好地表征及区分不同的微源投切位置.显然该方法可推广至 n -SMG 系统及类似采用 CPS-SPWM 调制的串联结构系统.为了进一步降低特征维数,定义工况 S_i 对应的三相微源投切中心欧式距离为

$$d_i = \sqrt{(\mu_a - \mu_b)^2 + (\mu_b - \mu_c)^2 + (\mu_c - \mu_a)^2} \quad (6)$$

式中: μ_j 的定义同式(4)。

对于一个投切工况 S_i , 通过上述四个特征即可获取其对应的投切微源数目和位置信息, 且特征维数从 $3n$ 个投切状态减少为 4 个, 这对提高模型训练效率有重要作用。

1.2 交互稳定性分级

基于阻抗的并网系统稳定性研究表明, 并网运行系统阻抗与电网阻抗之间的动态耦合决定了交互系统的稳定性。当两者阻抗幅值相交处相位裕度较小或过大时, 交互系统易出现阻抗交点频率附近的特定次谐波振荡, 从而导致交互系统不稳定, 而此时并网电流中相应谐波分量会因此而表现为增大至超过其限值^[14-15, 20-22]。SMSI-MG 中微源投切后, 其输出阻抗必然发生改变, 在假设运行环境其他参数不发生变化的情况下, 交互系统稳定性、并网电流谐波分量都会因此而发生改变。为避免复杂的序阻抗测量及相位裕度计算, 本文通过提取并网电流中的谐波分量衡量交互系统稳定性。同时, 为了充分利用机器学习处理多分类任务的优势, 不同于传统方法仅有稳定和 unstable 两个状态, 本文在稳定和 unstable 之间设置临界状态, 从而对交互系统稳定性进行更细致和直观的划分。

假设投切工况 S_i 时 PCC 处三相电流中谐波分量为 T_{hl} , 按照我国 GB/T 14549—1993 电能质量标准要求所得谐波电流分量的标准值为 T_{hl-sta} , 定义各次谐波超限比为 $\xi_{hl} = (T_{hl} - T_{hl-sta})/T_{hl-sta}$; 依据 ξ_{hl} 将交互系统稳定性划分等级如下

$$W_l = \begin{cases} 1 & \max_{h \in [2, 62]} \xi_{hl} \leq 0 \\ 2 & 0 < \max_{h \in [2, 25]} \xi_{hl} \leq \xi_{low1} \text{ 且 } \max_{h \in [26, 62]} \xi_{hl} \leq \xi_{hig1} \\ & \text{或 } \max_{h \in [2, 25]} \xi_{hl} \leq \xi_{low1} \text{ 且 } 0 < \max_{h \in [26, 62]} \xi_{hl} \leq \xi_{hig1} \\ 3 & \xi_{low1} < \max_{h \in [2, 25]} \xi_{hl} \leq \xi_{low2} \text{ 且 } \max_{h \in [26, 62]} \xi_{hl} \leq \xi_{hig2} \\ & \text{或 } \max_{h \in [2, 25]} \xi_{hl} \leq \xi_{low2} \text{ 且 } \xi_{hig1} < \max_{h \in [26, 62]} \xi_{hl} \leq \xi_{hig2} \\ 4 & \max_{h \in [2, 25]} \xi_{hl} > \xi_{low2} \text{ 或 } \max_{h \in [26, 62]} \xi_{hl} > \xi_{hig2} \end{cases} \quad (7)$$

式中: h 代表谐波阶次; ξ_{low1} 、 ξ_{low2} 为设定的 25 次以下谐波电流超限比阈值; ξ_{hig1} 、 ξ_{hig2} 为 25 次以上谐波电流超限比阈值; $W_l = 1$ 时谐波电流分量均在限值以下, 说明交互系统稳定裕度充足, 系统稳定, 该种投

切工况不会对系统安全稳定运行造成影响; $W_l = 2$ 表示较稳定, 此时特定次谐波电流分量有较小幅度的增加, 交互系统可能存在轻微谐振; $W_l = 3$ 时特定次谐波电流分量超过限值较多, 说明交互系统稳定裕度低, 系统进入不稳定状态。为与 $W_l = 4$ 区分, 称此时为较不稳定等级, 这时通过调整控制方式可对 S_i 引起的谐波振荡进行抑制; $W_l = 4$ 时特定次谐波电流含量超过 50%, 交互系统失稳, 且并网电流严重畸变, 在系统运行过程中应禁止出现该种投切工况。

为了有针对性地调整较不稳定情况下的控制策略, 及时抑制或消除微源投切引起的谐振, 对该谐振的频率范围进行预测是有必要的。对此, 将 $W_l = 3$ 的样本对应的谐波频段作为预测目标, 构建如下分类器。

$$HB_l = \begin{cases} 1 & \xi_{low1} < \max_{h \in [2, 25]} \xi_{hl} \leq \xi_{low2} \text{ 且 } \max_{h \in [26, 62]} \xi_{hl} \leq \xi_{hig1} \\ 2 & \xi_{hig1} < \max_{h \in [26, 62]} \xi_{hl} \leq \xi_{hig2} \text{ 且 } \max_{h \in [2, 25]} \xi_{hl} \leq \xi_{low1} \\ 3 & \xi_{low1} < \max_{h \in [2, 25]} \xi_{hl} \leq \xi_{low2} \text{ 且 } \xi_{hig1} < \max_{h \in [26, 62]} \xi_{hl} \leq \xi_{hig2} \end{cases} \quad (8)$$

式中, $HB_l = 1$ 代表造成交互系统不稳定的谐振发生在低频段; $HB_l = 2$ 则代表高频段; $HB_l = 3$ 代表既存在低频谐振也存在高频谐振。

2 用于交互稳定性评估的 GBDT 算法

2.1 GBDT 分类算法

梯度提升决策树^[23](GBDT)是一种基于决策树的集成学习方法, 其基本思想是每轮迭代中通过拟合残差来学习一个 CART 回归树模型 $T(x; \theta_m)$, 使得损失函数逐步减小。

对于具有 K 个类别的训练样本, GBDT 学习算法框架如图 3 所示。采用 One vs Rest (OvR) 分类方法将类别进行 one-hot 编码, 对每个编码输出分别建立决策树进行迭代。有监督的机器学习通常通过损失函数衡量预测值与真实值之间的偏差, 模型训练的目标是损失函数值尽可能小。多分类 GBDT 学习算法采用对数损失函数, 即

$$L(\{y_k, F_k(x)\}_1^K) = - \sum_{k=1}^K y_k \log p_k(x) \quad (9)$$

式中: $p_k(x) = \frac{e^{F_k(x)}}{\sum_{k=1}^K e^{F_k(x)}}$ 。文中训练样本集 $D = \{(x_l, y_l)\}$

($l = 1, 2, \dots, L; x_l \in \{\alpha_l, \beta_l, \gamma_l, d_l, \dots\}; y_l \in \{W_l, HB_l\}$)。

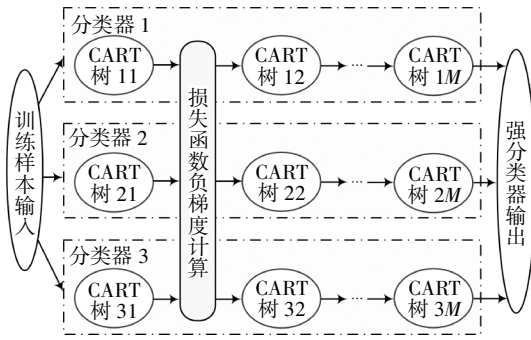


图3 GBDT 多分类模型框架

Fig.3 Frame of GBDT multi classification model

第 $t-1$ 轮迭代后,第 k 个分类器的负梯度误差为

$$r_{t-1,k}(x) = - \left[\frac{\partial L\{y_k, F_k(x)\}}{\partial F_k(x)} \right]_{\{F_k(x)=F_{t-1,k}(x)\}} = y_{ik} - p_{t-1,k}(x_i) \quad (10)$$

该误差是样本 l 对应类别 k 的真实标记和 $t-1$ 轮 GBDT 预测概率的差值,即残差.为使损失函数进一步减小,将 $(x_l, r_{t-1,k})$ 作为第 t 轮、第 k 个分类器中决策树的训练数据,如此迭代,直到模型准确度无法再提高时停止,获得最终的 GBDT 模型.

为防止过拟合,通过正则化项及子采样的方式提高模型泛化能力.其中,正则化项通过为每个弱学习器加上权重缩减系数 v ,弱化单个学习器的效果,从而提高模型整体的性能.加入正则化项后 GBDT 模型第 t 轮迭代输出为

$$F_t(x) = F_{t-1}(x) + vT(x; \theta_m) \quad (11)$$

另外,GBDT 模型中弱学习器之间存在依赖关系,其训练过程是串行进行的,但通过无放回抽样可实现部分并行训练,从而提高模型泛化能力及训练效率.

2.2 SMOTE 过采样

微源投切后交互系统中稳定和失稳样本较少,类别比例不平衡,这将导致分类结果准确率较低.对此,本文采用合成少数类过采样 (synthetic minority oversampling technique, SMOTE) 方法^[24]使原始数据中的类别比例更合理.该算法采用 k NN 方法在少数类样本的 k 个近邻中随机挑选 N 个样本进行随机线性插值,并与原数据合成产生新样本.

图4为部分样本采用 SMOTE 方法前后的类别分布对比图.其中,原始样本中 $W_l=1$ 的样本比例为 11.7%, $W_l=3$ 的样本占到 51%, $W_l=4$ 的样本则仅占 7.3%.经过 SMOTE 处理之后,原本占比较少的类别 1 和 3 的样本分别在其 k 近邻内通过插值产生了

与其类别一致的新样本,从而使得这两类样本数量明显增多,最终 3 类样本比例趋于平衡.

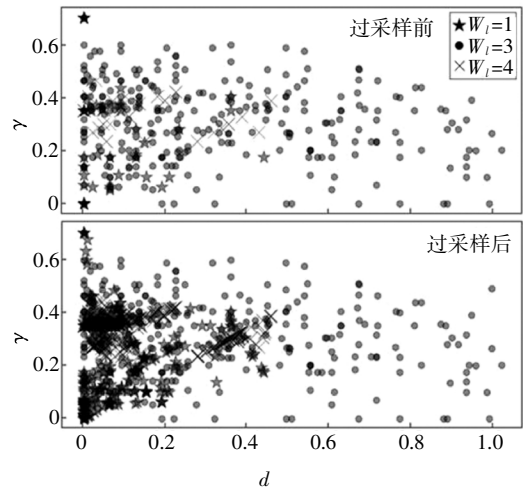


图4 SMOTE 前后样本类别分布对比

Fig.4 Comparison of sample category distribution before and after SMOTE

2.3 评价指标

交互稳定性分析中,若将较不稳定和失稳样本 ($W_l=3, W_l=4$) 错判为稳定,将导致交互系统失稳前运行人员不采取任何控制或禁止措施,从而给系统造成严重破坏.相比较而言,此种误判比将稳定样本误判为失稳样本带来的影响要大得多.因此,模型评价指标除考虑准确率 A_{cc} 外,还更加关注较不稳定和失稳样本的精度、召回率以及 F_1 值.多分类器混淆矩阵如表3所示.

表3 交互稳定性评估多分类器混淆矩阵

Tab.3 Confusion matrix of interactive stability evaluation classifiers

预测类别	实际类别			
	$W_l=1$	$W_l=2$	$W_l=3$	$W_l=4$
$W_l=1$	T_1	F_{12}	F_{13}	F_{14}
$W_l=2$	F_{21}	T_2	F_{23}	F_{24}
$W_l=3$	F_{31}	F_{32}	T_3	F_{34}
$W_l=4$	F_{41}	F_{42}	F_{43}	T_4

表3中, T_1, T_2, T_3, T_4 分别为稳定样本、较稳定样本、较不稳定样本和失稳样本被正确预测的数量; F_{pa} ($p=1, 2, 3, 4; a=1, 2, 3, 4$) 分别为预测为类别 p 而实际为类别 a 的错误分类的样本数量.

模型准确率表示正确分类的样本数占总体样本数的比值,即

$$A_{ccs} = \frac{T_1 + T_2 + T_3 + T_4}{T_1 + T_2 + T_3 + T_4 + F_{pa}} \quad (12)$$

较不稳定样本和失稳样本的精确度、召回率以及 F_1 值定义为

$$P_3 = \frac{T_3}{F_{31} + F_{32} + T_3 + T_{34}} \quad (13)$$

$$P_4 = \frac{T_4}{F_{41} + F_{42} + T_{43} + T_4} \quad (14)$$

$$R_3 = \frac{T_3}{F_{13} + F_{23} + T_3 + F_{43}} \quad (15)$$

$$R_4 = \frac{T_4}{F_{14} + F_{24} + F_{34} + T_4} \quad (16)$$

$$F_{1-i} = \frac{2P_i R_i}{P_i + R_i}, i = 3, 4 \quad (17)$$

较不稳定样本中的谐振频段预测混淆矩阵如表 4 所示。

表 4 谐振频段预测多分类器混淆矩阵

Tab.4 Confusion matrix of resonant frequency band predicton classifiers

预测类别	实际类别		
	HB _i = 1	HB _i = 2	HB _i = 3
HB _i = 1	T_1	F_{12}	F_{13}
HB _i = 2	F_{21}	T_2	F_{23}
HB _i = 3	F_{31}	F_{32}	T_3

定义模型谐振频段预测准确率为

$$A_{\text{cch}} = \frac{T_1 + T_2 + T_3}{T_1 + T_2 + T_3 + F_{pa}} \quad (18)$$

3 仿真分析

3.1 样本生成

通过对 10-SMG 系统中各种预想投切工况进行大量仿真,获取样本数据。仿真中,为保证投切前后并网运行电流方向不变,取 $\alpha_i \leq 43.3\%$ 。另外,以初始状态为基准,设定电网阻抗和滤波电感变化范围为 $\pm 25\%$ 。控制器采用准比例谐振,其比例系数变化范围为 0.01~0.5。谐波电流超限比阈值设定为: $\xi_{\text{low}1} = 1.2, \xi_{\text{low}2} = 8, \xi_{\text{high}1} = 1.5, \xi_{\text{high}2} = 45$ 。经过筛选后获取 1 000 组样本数据,并以其中 70% 作为训练样本集,30% 作为验证样本集。

为保证测试样本的均匀性,利用线性变换对属性值进行归一化处理

$$X_{\text{sca}} = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \quad (19)$$

式中, $X_{\min}$ 和 $X_{\max}$ 分别为某特征属性的最小和最大值。

3.2 不同模型仿真结果对比

为验证模型评估和预测效果,将 GBDT 模型与其他常用机器学习模型的评估结果进行对比。测试

模型均基于开源机器学习库 Scikit-learn 搭建^[25]。参数使用五折交叉验证的方法获取最优值,GBDT 模型的正则化通过超参数 $n_estimators$ 、 $learning_rate$ 和 $subsample$ 实现;支持向量机(SVM)使用多项式核函数。采用式(12)~(18)定义的评价指标,得到验证集上的测试情况如表 5 所示。其中,GBDT 表示使用未经 SMOTE 处理的训练数据得到的模型,除此之外的模型均采用过采样后的样本数据;谐振频段预测则使用原始样本中 $W_i = 3$ 的样本。

表 5 不同模型评估和预测结果对比

Tab.5 Comparison of assessment and prediction results of different models

模型	$A_{\text{cch}}/\%$	F_{1-3} 值/ $\%$	F_{1-4} 值/ $\%$	$A_{\text{cch}}/\%$
GBDT+SMOTE	98.37	98.49	100	—
GBDT	92.09	90.09	100	98.65
RF	98.37	95.86	100	97.04
TREE	97.66	98.32	99.5	96.69
SVM	96.49	96.61	100	97.04
kNN	94.73	92.02	99.5	95.09

表 5 中,GBDT 和 RF 为集成学习器,其余为个体学习器。对比发现,集成学习器总体比个体学习器拥有较高的准确率;同样是集成学习器,RF 和 GBDT 准确率相同,但 GBDT 在类别 3 上具有较高的 F_1 值,且谐振频段预测准确率较高。未经过采样的 GBDT 模型评估准确率最低,可见样本类别平衡对提高模型准确率有重要作用,由此也证明了结合过采样方法的必要性和有效性。总体看来,结合 SMOTE 的 GBDT 模型可实现交互系统稳定性和谐振频段较高准确率的评估和预测,这也从另一个方面说明了特征提取方法的正确性。

3.3 模型迁移能力验证

为验证所建模型的适应性,设置微源数目 n 从 5 至 40 以 5 为步长变化,分别建立 n -SMG 并网运行系统仿真模型,且新拓扑系统保持与 10-SMG 直流侧总电压相等。各系统分别对 30 种不同投切 S_i 进行仿真构成 240 个新建测试样本。由于微源数目 n 不同,在利用式(5)、(6)计算投切微源位置特征属性值时,会出现与 10-SMG 系统规律相同但属性值不同的情况,使得之前构建的 GBDT 模型应用于 n -SMG ($n \neq 10$) 系统时会发生误判。因此对于 n -SMG ($n \neq 10$) 系统,首先需要对特征 γ 和 d 进行映射,使其与 10-SMG 的数值相对应。

选取若干具有代表性的点对 $[x_{n-SMG}, x_{10-SMG}]$,通过线性拟合获得新拓扑系统和 10-SMG 在特征 X 下的对应关系 $g(X)$,如图 5 所示的 20-SMG 与 10-SMG 系统中的 $g(\gamma)$ 和 $g(d)$. 之后将 $g(X)$ 应用于新拓扑下的其他样本,即可获得映射后的特征属性值.通过该映射方法,分别得到 n -SMG 修正后的特征值,通过之前所建 GBDT 模型进行交互稳定性分级评估和谐振频段预测,验证结果如表 6 所示.其中 10-SMG 的投切样本是原始样本中未出现的投切工况.

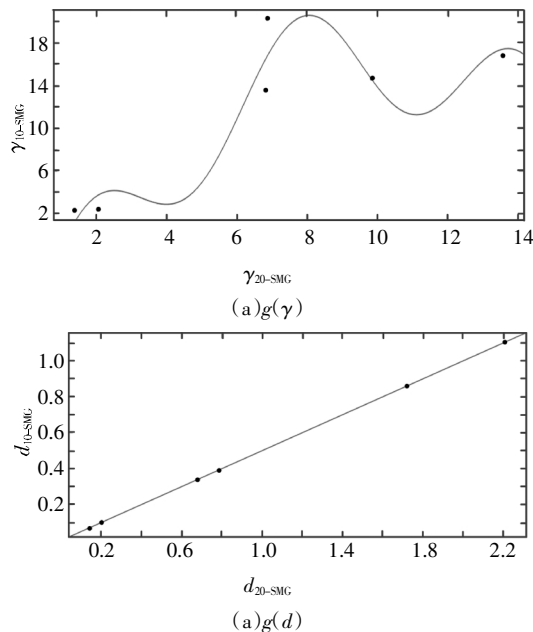


图 5 20-SMG 与 10-SMG 中 γ 和 d 的映射关系

Fig.5 The mapping relationship between γ and d in 20-SMG and 10-SMG systems

表 6 模型迁移能力验证

Tab.6 Verification results of model migration capability

	$A_{ces}/\%$	$A_{ech}/\%$	F_{1-3} 值/ $\%$	F_{1-4} 值/ $\%$
5-SMG	100	100	100	100
10-SMG	100	100	100	100
15-SMG	96	95	96.6	90
20-SMG	96	100	96.3	100
25-SMG	95	95	96.3	92
30-SMG	96.27	98	95.3	100
35-SMG	93	93.7	94.6	90
40-SMG	94	94.6	95	98

从表 6 可看出,所建模型应用于 5-SMG 和 10-SMG 时,测试准确度较高.而应用于 $n>10$ 的系统时,由于 γ 的非线性映射以及训练样本的缺失,准确率

有所下降,但后期通过补充样本可望进一步提高准确率.总体来看,所建模型通过特征属性的简单拟合即可适用于新拓扑系统,避免花费大量时间重新生成样本,具有一定的实际应用价值.

基于 GBDT 的交互稳定性评估主要由离线训练和在线评估两部分构成.通过前期数据集生成、离线模型训练和模型性能评估,可获得训练好的 GBDT 模型.在线评估时通过在线获取微源投切等特征数据,并将特征数据输入(根据系统拓扑结构,映射或直接输入)训练好的模型,即可对相应投切操作引起的交互稳定性进行评估,并预测谐振频段,为预防控制提供及时有效的决策依据.

4 结 论

本文将机器学习用于微源投切后的 SMSI-MG 与电网交互稳定性分析,提出一种适用于串联结构系统的微源投切特征提取方法,并采用 GBDT 模型进行交互稳定性评估和谐振频段预测.得到如下结论:

1)对于采用 CPS-SPWM 调制的并网运行 SMSI-MG 系统,其与电网的交互稳定性不仅与投切微源数目相关,还与投切微源的位置相关.通过三相微源平均投切率、投切不平衡度、投切平均离散度、投切中心欧式距离四个特征函数即可提取三相投切微源的数目和位置信息.而通过对后两者的拟合,模型即可具备良好的泛化和迁移能力.特征提取方法正确,模型训练效率高.

2)将微源投切对交互系统稳定性的影响划分为 4 个类别,类别划分更细致,且对较不稳定样本中的谐振频率范围进行预测,避免了传统方法中的谐振检测.模型输出为多类别、多标签,对后续决策更具实时性和指导性.

3)采用结合过采样的集成学习方法,与其他个体学习器的评估方法相比,保证了较高的总体评估和预测准确性.模型泛化能力强,具有实用价值.

本文所提基于机器学习模型的交互稳定性评估方法无需建立复杂的数学模型,为并网运行系统稳定性在线分析及决策提供了新思路,同时所提出的特征提取方法也可为类似串联结构系统在电能质量分析、优化运行等方面提供方法参考.在今后的研究中,除进一步扩大样本规模外,还需寻找更优的指标将微源投切对交互系统的影响量化呈现出来,以进一步提高模型的实际应用价值.

参考文献

- [1] YANG W M, WANG X G, LI X Y, *et al.* An active power sharing method among distributed energy sources in an islanded series micro-grid[J]. *Energies*, 2014, 7(12): 7878—7892.
- [2] 杨维满, 王兴贵. 串联结构中H桥逆变器直流侧等效模型关系[J]. *电工技术学报*, 2015, 30(16): 103—110.
YANG W M, WANG X G. Relationship of DC-side equivalent model among H-bridge inverters connected in series[J]. *Transactions of China Electrotechnical Society*, 2015, 30(16): 103—110. (In Chinese)
- [3] 王兴贵, 杨维满. 孤岛模式下串联型微电网逆变系统动态建模[J]. *高电压技术*, 2014, 40(8): 2456—2463.
WANG X G, YANG W M. Dynamic modeling of inverter system in isolated series micro-grids[J]. *High Voltage Engineering*, 2014, 40(8): 2456—2463. (In Chinese)
- [4] XIAO B L, HANG L J, MEI J, *et al.* Modular cascaded H-bridge multilevel PV inverter with distributed MPPT for grid-connected applications [J]. *IEEE Transactions on Industry Applications*, 2015, 51(2): 1722—1731.
- [5] GE B M, LIU Y S, ABU-RUB H, *et al.* State-of-charge balancing control for a battery-energy-stored quasi-Z-source cascaded-multilevel-inverter-based photovoltaic power system [J]. *IEEE Transactions on Industrial Electronics*, 2015, 65(3): 2268—2279.
- [6] SHUAI Z K, LIU D G, SHEN J, *et al.* Series and parallel resonance problem of wideband frequency harmonic and its elimination strategy[J]. *IEEE Transactions on Power Electronics*, 2014, 29(4): 1941—1952.
- [7] 刘怀远, 徐殿国, 武健, 等. 并网换流器系统谐振的分析、检测与消除[J]. *中国电机工程学报*, 2016, 36(4): 1061—1074.
LIU H Y, XU D G, WU J, *et al.* Analysis, detection and mitigation of resonance in grid-connected converter systems [J]. *Proceedings of the CSEE*, 2016, 36(4): 1061—1074. (In Chinese)
- [8] WANG X F, BLAABJERG F, WU W M. Modeling and analysis of harmonic stability in an AC power-electronics-based power system [J]. *IEEE Transactions on Power Electronics*, 2014, 29(12): 6421—6432.
- [9] WANG X F, BLAABJERG F, LOH P C. Grid-current-feedback active damping for LCL resonance in grid-connected voltage-source converters [J]. *IEEE Transactions on Power Electronics*, 2016, 31(1): 213—223.
- [10] 朱晓娟, 胡海涛, 陶海东, 等. 光伏并网系统的谐波不稳定产生机理及影响规律[J]. *电工技术学报*, 2017, 32(10): 33—41.
ZHU X J, HU H T, TAO H D, *et al.* The mechanism and influence of harmonic instability for photovoltaic grid-connected system [J]. *Transactions of China Electrotechnical Society*, 2017, 32(10): 33—41. (In Chinese)
- [11] HARNEFORS L, WANG X F, YEPES A G, *et al.* Passivity-based stability assessment of grid-connected VSCs—an overview[J]. *IEEE Journal of Emerging and Selected Topics in Power Electronics*, 2016, 4(1): 116—125.
- [12] KWON J, WANG X F, BAK C L, *et al.* Harmonic interaction analysis in grid connected converter using Harmonic State Space (HSS) modeling [C]//2015 IEEE Applied Power Electronics Conference and Exposition (APEC). Charlotte, NC, USA: IEEE, 2015: 1779—1786.
- [13] BOTTRELL N, PRODANOVIC M, GREEN T C. Dynamic stability of a microgrid with an active load [J]. *IEEE Transactions on Power Electronics*, 2013, 28(11): 5107—5119.
- [14] SUN J. Impedance-based stability criterion for grid-connected inverters[J]. *IEEE Transactions on Power Electronics*, 2011, 26(11): 3075—3078.
- [15] 陈新, 张旸, 王赞程. 基于阻抗分析法研究光伏并网逆变器与电网的动态交互影响 [J]. *中国电机工程学报*, 2014, 34(27): 4559—4567.
CHEN X, ZHANG Y, WANG Y C. A study of dynamic interaction between PV grid-connected inverters and grid based on the impedance analysis method [J]. *Proceedings of the CSEE*, 2014, 34(27): 4559—4567. (In Chinese)
- [16] 石访, 张林林, 胡熊伟, 等. 基于多属性决策树的电网暂态稳定规则提取方法[J]. *电工技术学报*, 2019, 34(11): 2364—2374.
SHI F, ZHANG L L, HU X W, *et al.* Power system transient stability rules extraction based on multi-attribute decision tree[J]. *Transactions of China Electrotechnical Society*, 2019, 34(11): 2364—2374. (In Chinese)
- [17] 周挺, 杨军, 周强明, 等. 基于改进 LightGBM 的电力系统暂态稳定评估方法[J]. *电网技术*, 2019, 43(6): 1931—1940.
ZHOU T, YANG J, ZHOU Q M, *et al.* Power system transient stability assessment method based on modified LightGBM[J]. *Power System Technology*, 2019, 43(6): 1931—1940. (In Chinese)
- [18] 张晨宇, 王慧芳, 叶晓君. 基于 XGBoost 算法的电力系统暂态稳定评估[J]. *电力自动化设备*, 2019, 39(3): 77—83.
ZHANG C Y, WANG H F, YE X J. Transient stability assessment of power system based on XGBoost algorithm [J]. *Electric Power Automation Equipment*, 2019, 39(3): 77—83. (In Chinese)
- [19] 周京华, 陈亚爱. 高性能级联型多电平变换器原理及应用[M]. 北京: 机械工业出版社, 2013: 87.
ZHOU J H, CHEN Y A. Principle and application of high performance cascaded multilevel converter [M]. Beijing: China Machine Press, 2013: 87. (In Chinese)
- [20] CESPEDAS M, SUN J. Impedance modeling and analysis of grid-connected voltage-source converters [J]. *IEEE Transactions on Power Electronics*, 2014, 29(3): 1254—1261.
- [21] 王赞程, 陈新, 陈杰, 等. 基于谐波线性化的三相 LCL 型并网逆变器正负序阻抗建模分析 [J]. *中国电机工程学报*, 2016, 36(21): 5890—5898.
WANG Y C, CHEN X, CHEN J, *et al.* Analysis of positive-sequence and negative-sequence impedance modeling of three-phase LCL-type grid-connected inverters based on harmonic linearization [J]. *Proceedings of the CSEE*, 2016, 36(21): 5890—5898. (In Chinese)
- [22] CHO Y, HUR K, KANG Y C, *et al.* Impedance-based stability analysis in grid interconnection impact study owing to the increased adoption of converter-interfaced generators [J]. *Energies*, 2017, 10(9): 1355—1372.
- [23] FRIEDMAN J H. Greedy function approximation: a gradient boosting machine[J]. *Annals of Statistics*, 2001, 29(5): 1189—1232.
- [24] CHAWLA N V, BOWYER K W, HALL L O, *et al.* SMOTE: synthetic minority over-sampling technique [J]. *Journal of Artificial Intelligence Research*, 2002, 16: 321—357.
- [25] PEDREGOSA F, VAROQUAUX G, GRAMFORT A, *et al.* Scikit-learn: Machine learning in Python [J]. *Journal of Machine Learning Research*, 2011, 12: 2825—2830.