

多输入特征融合的组合支持向量机 电力系统暂态稳定评估

马 骞¹, 杨以涵¹, 刘文颖¹, 齐 郑¹, 郭金智²

(1. 华北电力大学电力工程系, 北京市 昌平区 102206; 2. 华北电力调度局, 北京市 宣武区 100053)

POWER SYSTEM TRANSIENT STABILITY ASSESSMENT WITH COMBINED SVM METHOD MIXING MULTIPLE INPUT FEATURES

MA Qian¹, YANG Yi-han¹, LIU Wen-ying¹, Qi Zheng¹, Guo Jin-zhi²

(1. Department of Electrical Power Engineering, North China Electric Power University, Changping District, Beijing 102206, China; 2. North China Electric Power Dispatch Bureau, Xuanwu District, Beijing 100053, China)

ABSTRACT: In the assessment of the transient stability with the Support Vector Machine (SVM), the choices of the input features are the most important factors to the final results. The traditional heuristic methods and the tentative methods cannot solve this problems radically. This paper, with the idea of information fusion, proposes a two-layer SVM classifier model to improve the precision of the classification. In this paper, four different groups of input features are built based on the heuristic knowledge of different angles to form four weak classifiers. By taking these four weak classifiers as the sub-classifiers, a fusional SVM is built to fuse the results of the sub-classifiers in a manner of regression, whose result is taken as the final judgment result. Simulation results on the IEEE 39-BUS and IEEE 145BUS test system show that the classification performance of the weak classifiers is evidently strengthened after fusion, and that the result after the fusion is more precise than any of the single results and the result including all the input features at once. The proposed method is a promising tool of fast computation for on-line transient stability assessment.

KEY WORDS: Power system; Transient stability assessment; Support Vector Machine (SVM); Feature selection; Information fusion

摘要: 利用支持向量机(SVM)方法进行暂态稳定判别时, 输入特征的选择是影响最终结果的最重要因素。传统启发式和试探式方法不能从根本上解决输入特征选择的问题。本文利用信息融合思想, 在构造的具有不同输入特征的多组子分类器的基础上, 对子分类器的结果在输出空间再进行信息融合, 以提高分类准确率。文中从不同角度启发式的构造了 4

组不同的输入特征, 构成四组弱分类器。以这四组弱分类器为子分类器, 再构造一个融合 SVM 对几种子分类器的结果以回归方式进行融合, 作为最终判别结果。IEEE 39-BUS 和 IEEE145-BUS 测试系统上进行的仿真表明, 弱分类器的分类性能经过融合得到明显强化, 融合后的结果比任何一种子分类器的结果以及一次包含所有输入特征的结果都更准确。该方法为在线快速进行暂态稳定计算提供了一条重要途径。

关键词: 电力系统; 暂态稳定判别; 支持向量机(SVM); 特征选择; 信息融合

1 引言

利用人工智能方法进行非模型的电力系统暂态稳定判别(Transient Stability Assessment TSA)具有在线计算速度快、容易生成决策用的启发规则等优点, 可与传统暂稳分析方法构成良好的互补^[1-3]。并且, 近年来监测系统在线运行积累了大量的数据, 利用这些数据所提供的各种运行方式下的预想事故分析结果, 可以进行有意义的机器学习, 以便从大量的数据特征中找到进行暂态稳定评估的有效方法。

以支持向量机(Support Vector Machine SVM)为代表的机器学习方法近年来得到飞速发展并在电力系统中取得广泛应用^[4-8]。相比较于已有的各种学习方法, 支持向量机方法的优势在于:

(1) 结构风险最小保证了“最坏情形下”的学习机器的泛化能力;

(2) 其凸二次优化问题保证了惟一解, 从而避

免了神经网络遇到的局部极小问题;

(3) 能有效地对小样本进行处理(实际上, 无论所采用的训练样本有多丰富, 相比于原问题都是小样本)。

但单纯的 SVM 也有其缺点。就应用于模式分类的 SVM 而言, SVM 最大的应用障碍在于自身无法确定最合理的输入特征组合^[9]。理想的输入特征组合是能将样本分为只包含正例或只包含反例的集合^[9]。而在实际问题中由于几乎无法获得关于实际问题的全部事实而使最优特征组合无法被获得^[9]。其次, SVM 的核参数选择至今无通用方法。典型常用的核函数不一定是针对给定问题的最好映射^[10], 映射后的高维空间也不一定是最高效的分类空间。

对输入特征选择困难的问题, 本文提出了一种基于信息融合的组合 SVM 模型, 并将之应用于解决电力系统暂态稳定判别问题。仿真表明, 组合 SVM 的引入很好地解决了输入特征选择困难的问题, 显著地提高了最终判别结果的准确率。

2 统计学习与支持向量机简介

2.1 学习问题的一般描述

从样本进行学习的一般见图 1 所示^[11]。

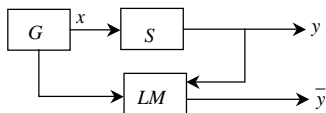


图 1 学习问题的一般模型
Fig. 1 General modal of learning problem

学习机器 LM 观察来自产生器 G 的 x 和训练器 S 的 y 所构成的数据对 (x, y) , 从给定的函数集中选择出能最好逼近训练器响应的函数。在训练之后, 学习机器必须对任意输入 x 给出输出 \bar{y} 。学习的目标是能够给出尽可能接近训练器响应 y 的输出 \bar{y} 。当采用完全有界非负函数集的学习机器时, 其推广能力的界为^[11]

$$R(\alpha_l) \leq R_{emp}(\alpha_l) + \frac{B\varepsilon}{2} \left(1 + \sqrt{1 + \frac{4R_{emp}(\alpha_l)}{B\varepsilon}} \right) \quad (1)$$

式(1)不等式左边为总风险; 右端第一项为经验风险, B 、 ε 均为与函数集有关的参数。

由式(1)可见, 总风险不仅与经验风险有关, 还与学习机器采用的函数集的性质有关。为使总风险最小, 还必须使学习机器所采用的函数集满足使第 2 项达到最小, 即所谓的结构风险最小原则。SVM 正是求解学习机器函数集的结构风险泛函最小化的

体现。下面通过分类问题介绍 SVM 的一些概念。

2.2 支持向量机

简单地考虑分类问题时, 设有一组样本 $(x_1, y_1), (x_2, y_2), \dots, (x_l, y_l)$, 其中 $x \in R^n$, 为输入, $y \in \{-1, 1\}$, 为输出, 分类器的学习问题描述如下:

当样本为线性可分时, 训练数据可被超平面没有错误地分开为

$$(w * x) - b = 0 \quad (2)$$

并且离超平面最近的向量与超平面之间的距离最大时, 则说这个超平面是这个向量集合的最优分类超平面。最优分类超平面是结构风险最小化的结果。距离最优超平面最近的两类样本向量称为支持向量。

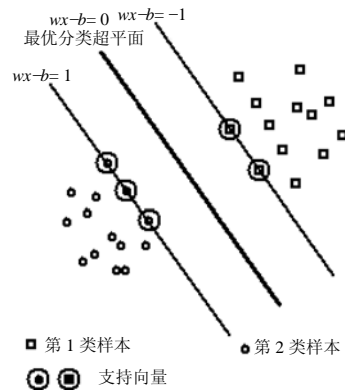


图 2 最优超平面示意图

Fig. 2 Demonstration figure of optimal hyper-plane for classification

在线性可分的情况下, 由于两类支持向量到最优超平面间的距离之和为 $2/\|w\|$, 故构造最优超平面的问题被转化为如下的优化问题:

$$\begin{cases} \min \Phi(w) = \|w\|^2 = (w \times w) / 2 \\ \text{st } y_i(w \times x_i - b) \geq 1, i = 1, 2, \dots, l \end{cases} \quad (3)$$

该不等式约束下的凸二次优化问题存在惟一解为

$$w_0 = \sum_{i=1}^l \alpha_i^0 y_i x_i \quad (4)$$

式中 对所有 $i = 1, \dots, l, \alpha_i^0 \geq 0, \alpha_i$ 为 Lagrange 乘子。

在多数情况下, 式(4)的 w_0 展开式的 α_i^0 均为 0, 而非零的 α_i^0 所对应的 x_i 即为支持向量 (SV)。

通常训练集中的数据是线性不可分的, 这时需要将原问题映射到某个高维空间(可能是无穷大维数空间)求解, 样本在此高维空间内成为线性可分的。根据泛函理论, 只要一种核函数 $K(x_i, x_j)$ 满足 Mercer 条件, 它就对应某一变换空间中的内积。所以在最优分类面中采用适当的内积函数, 就可以对

样本实现某一非线性变换后的线性分类。此时最优分类判别函数为

$$f(x) = \text{sgn}\left(\sum_{SVS} y_i \alpha_i^0 K(x_i * x) - b_0\right) \quad (5)$$

式(5)即为支持向量机的分类判别函数。

3 组合 SVM 解决暂态稳定判别问题

3.1 现有 SVM 方法解决 TSA 问题的步骤

暂态稳定分析可以抽象为根据特征进行模式分类的问题，以利于用机器学习实现。现有的利用 SVM 方法解决 TSA 问题的步骤为：

(1) 训练集和测试集的准备。训练集和测试集要求独立同分布，训练集和测试集要同时构造，随机抽取，故一般做法是针对特定的网络进行大量运行方式下的独立故障仿真，随机抽取一部分作为训练样本，以保证样本的独立同分布要求，另一部分则作为测试样本。

(2) 选择合适的输入特征。传统的做法是根据经验从大量的描述实际问题的特征中选择若干作为 SVM 的输入，对 TSA 问题，常选用的特征为发电机故障切除瞬间的相角、转速、加速功率等。

(3) 选择 SVM 的核函数、参数。通常选择的核函数有高斯型、多项式型和感知器型，每种核函数都有参数需要确定。参数的确定可以用 CV(Cross Validation) 算法寻优^[12]。

(4) 利用训练集对 SVM 进行训练，得到分类判别函数 $f(x)$ 。

(5) 将训练完的分类器在测试集上测试。

采用常规 SVM 方法进行暂稳评估的最大问题就是输入特征无法准确选择。如果输入特征选择太少，则无法全面表征暂态稳定问题，得不到良好的训练效果；如果输入特征选择太多，相当于引入了与稳定无关的噪声，也得不到好的学习性能。理论证明指出：能表征问题的最少特征量的方法是最优的方法^[13]。已有的应用于暂态稳定判别的 SVM 方法研究中，对输入特征的选择问题尚处于启发、试探阶段。在文[7,8]中，均阐述了利用 SVM 进行暂态稳定判别的问题，虽然都得到了很好的分类效果，但均未对输入特征的选择问题给予充分的论证。

3.2 基于信息融合的组合 SVM 模型

对输入特征选择困难的问题，常用的方法为利用遗传算法对大量输入特征进行逐一选择，但该算法效率低，工作量大。对此利用信息融合思想，提出了一种组合 SVM 分类器模型：首先设法构造多

组输入特征，分别训练，得到相对而言分类效果较弱的几组 SVM 子分类器；再对这些 SVM 子分类器的输出结果做信息融合，采用融合后的输出作为最终的分类结果，如图 3 所示。

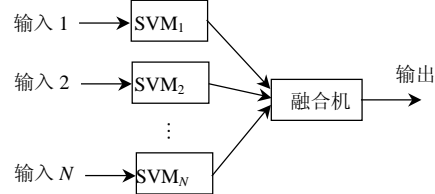


图 3 SVM 的信息融合模型图
Fig. 3 Demonstration figure of information fusion for SVM

信息融合的方法有很多种，由于 SVM 能提供结构风险最小的学习结果，故融合机仍使用了 SVM 二分类器。具体为：将第 1 层子分类器的输出以回归的形式作为另一个融合 SVM（第 2 层）的输入进行训练，即将第 1 层子分类器的实值输出直接作为第 2 层子分类器的输入值，以融合 SVM 的输出作为最终的分类结果。使用回归形式连接 2 层 SVM 的理由有：①第 1 层子分类器输出的实值具有表征当前分类结果准确性的意义，一般而言，所输出的实值越大，说明新样本距离分类超平面越远，分类结果越可信；②如果直接使用仅含有 0、1 二值的向量作为第 2 层 SVM 的输入，将很难保证输出具有好的泛化能力。为了使第 1 层的不同子分类器的输出具有同等数量级，在本文的模型中，将第 1 层子分类器输出的实值做了归一化处理如下：

$$I'_i = \frac{I_i}{D_i}, \quad i=1,2,3,4 \quad (6)$$

式中 I'_i 为归一化后的输出； I_i 为归一化前的实值输出； D_i 为该子分类器的支持向量到最优超平面的距离。

从形式上看，2 层的组合 SVM 分类器模型只是简单地将几个独立的 SVM 做连接，而这种连接背后却有其实际意义。可将第 1 层的若干个 SVM 分类器看成子分类器，每个子分类器独立工作，给出判别结果。第 2 层的 SVM 分类器则是对第 1 层的若干个子分类器的分类结果做综合评判，将最可靠的结果作为最终输出结果，以提高分类精度。

使用 2 层 SVM 信息融合模型进行 TSA 问题的解决过程如下：

(1) 训练集、测试集的准备。将训练集再分为 2 部分，第 1 部分用于训练第 1 层的 SVM 子分类器，第 2 部分用于训练融合 SVM；

(2) 输入特征的构造。需要根据不同角度对实际问题的分析独立选择多组输入,作为第1层 SVM 的输入;

(3) 利用训练集的第1部分训练第1层上的 SVM 子分类器;

(4) 将第1层 SVM 子分类器的输出以回归的形式作为第2层融合 SVM 的输入,在训练集的第2部分上进行训练,得到融合层 SVM;

(5) 融合后的 SVM 在测试集上进行测试。

4 多组输入特征的构造

4.1 输入特征分析

本文根据不同的启发知识选择了4组输入特征,作为第1层上的若干 SVM 子分类器的输入。其中的3组输入特征分别是基于以下几种对暂态稳定问题的分析得到的:基于超平面与注入功率空间、基于 CUEP 处的临界势能、基于功角曲线的 Taylor 展开式^[14]。另一组是已被证明具有良好的 ANN 学习能力的输入^[15]。这4组特征由于仅凭启发式分析得到,故所构造的分类器均属于弱分类器,但经过信息融合,可以得到这些弱分类器融合后形成的强分类器,后面的仿真充分说明了这一点。

4.2 基于超平面和功率注入空间选择特征

已经证明,在一定近似条件下,电力系统的稳定边界在 UEP 点处可以用超平面近似^[16]。超平面方程如下^[17]:

$$\mathbf{x}^T \boldsymbol{\eta}_1 = \mathbf{x}_u^T \boldsymbol{\eta}_1$$

其中, $\mathbf{x}^T = (\boldsymbol{\alpha}^T, \boldsymbol{\omega}_g^T)$, 为超平面自由变量; $\boldsymbol{\eta}_1$ 为发电机状态方程在 UEP 处线性化后所对应的雅克比矩阵 \mathbf{J}_{x^T} 的正特征值的特征向量; \mathbf{x}_u 为超平面处的 UEP。

实用的超平面方程为^[18]

$$(\boldsymbol{\eta}^T \mathbf{S}_2)^{-1} \boldsymbol{\eta}^T \mathbf{S}_1 \mathbf{P} = \sum_{i=1}^r \alpha_i P_i = 1. \quad (7)$$

式中 P_i 为节点注入功率; $\boldsymbol{\eta}$ 为超平面的法矢量, $\mathbf{S}_1, \mathbf{S}_2$ 为符合下列等式的矩阵^[18]:

$$\begin{bmatrix} \dot{\alpha} \\ \dot{\omega}_g \end{bmatrix}_r = \mathbf{S}_1 \mathbf{P}' + \mathbf{S}_2$$

以上超平面方程的许多参数需要进行大量计算才能得到(某些参数与 CUEP 有关,准确求取 CUEP 需耗费大量计算),只能间接选择与这些参数相关度较大的且方便获得的一些参数。故选择基于超平面

启发的输入特征如下:

(1) 选择故障后关键节点的注入功率 P_i 。因为关键节点原指与骨干网相连的节点、区域代表性节点和需要观察的节点^[16],在本文中,关键节点选择为节点注入功率数值超过系统平均节点注入功率(系统总容量除以系统中电源节点数)的节点;

(2) 选择故障切除瞬间前绝对转速最大的 k 个和最小的 k 个 δ_i, ω_i 。因为绝对转速最大的和最小的发电机节点影响 CUEP 相对于 SEP 的位置;

(3) 选择故障后线路两端相角差最大的前 k 个线路的 P_{ij} 和 δ_{ij} 。因为相角差最大的线路通常属于临界割集;

(4) 选择在故障前、后 SEP 电压幅值变化最大的 k 个节点电压 U_i 。因为超平面求解中采用了事故期间各节点电压值保持不变的假设。

应指出的是:选择的输入特征(1)~(4)中的 k 不必均取相同数值,并且不宜取得过大,因为过多的特征反而会带来噪声因素。

4.3 基于功角曲线展开选择特征

将功角曲线进行 Taylor 级数展开,可以得到近似的故障后最大功角差的表达式^[14],基于此,文[14]共选择了13组输入特征作为 ANN 的输入特征。本文采用这些特征作为 SVM 的输入。

4.4 类相比于已有 ANN 选择特征

文[15]给出了已被验证具有良好学习能力的一组 ANN 的输入特征,并分别给出了相关性系数。本文将这组特征也作为一个 SVM 子分类器的输入。从后面仿真可以看出,虽然 SVM 理论上比 ANN 更好的学习和泛化能力,但由于核函数及参数的选择等原因,在同样输入条件下并不能保证所得到的结果总好于 ANN。这一点在美国邮政的数字识别试验中也有所体现^[13]。

4.5 基于 CUEP 处临界势能值选择特征

在忽略暂态过程中的非自治因素条件下, CUEP 处的临界势能函数值可以作为判断系统稳定的一个依据。据此,可以构造基于 CUEP 处临界势能值的输入特征。求取 CUEP 处临界势能的具体求解过程见文[16]。根据文[16]给出的结论,基于 CUEP 处的临界势能方法启发的输入特征选取如下:

(1) 故障支路的电纳 B_{ij} 及故障前的有功 P_{ij} ;

(2) 故障后两端相角差最大的 k 个线路的 B_{ij} 、转速差最大的 k 个线路的 B_{ij} ;

(3)最大的前 k 个故障瞬间节点注入功率失配量 ΔP_i ;

(4)故障切除时刻最大的前 k 个 δ_i 和较小的 k 个 δ_i ;

(5)故障最近的发电机节点的 P_{mi} 、负荷节点的 P_{Li} 和关键线路的 P_{ij} 。关键线路是指故障前、后功率变化最大的(受冲击最大的线路);

(6)故障切除时刻最大和最小的 k 个 ω_i ;

(7)故障切除时刻最大的 k 个发电机动能 W_{ki} ;

(8)最大的发电机内电势 E_i ，及受冲击最大的前 k 个发电机的 E_i 。

5 仿真分析

本文在 IEEE39-BUS 系统和 IEEE145-BUS 系统上进行仿真。对于 IEEE39-BUS 系统,发电机模型采用经典模型,选取故障为三相短路,故障清除时间为 0.1s。在 70%~130%基准负荷范围内随机设置负荷并配置相应的发电出力,随机设置不同的故障位置及故障元件,共得到 600 个样本。随机抽取其中 400 个样本作为训练集(其中,200 个样本用来训练第 1 层子 SVM,200 个用来训练第 2 层 SVM),其余 200 个作为测试集。在 IEEE145-BUS 系统中,则增加了综合负荷模型,在 B₃₄、B₃₅、B₅₁、B₆₆、B₇₄、B₉₃ 等几个负荷较重的节点上使用了综合负荷模型,负荷均由马达负荷、恒功率负荷、恒阻抗负荷及恒电流负荷组成,组成比例由程序随机生成。其中的马达负荷构成比例<30%,恒功率<30%,恒电流<30%,剩余为恒阻抗。共得到 800 个样本。随机抽取其中 600 个作为训练集(其中,300 个样本用来训练第 1 层子 SVM,300 个用来训练第 2 层 SVM),其余 200 个作为测试集。

在实际系统中,大量的实验样本同样只能靠仿真获得,因为暂态稳定的实测、实验数据是难以获得的。SVM 对训练与测试样本的要求是符合独立同分布规律(即所有测试样本与训练样本间要服从同一分布函数)且各组样本间彼此是独立的。因此,基于实际系统中已有的运行方式数据,所进行的预想故障计算结果即可构成 SVM 的训练、测试样本集。从这些样本中随机抽取一部分作为训练样本集,另一部分作为测试样本集,即可满足独立同分布要求。值得指出的是,这些计算结果并不需要单独为 SVM 计算准备。事实上,所有电力公司的相关部门为了制定系统运行方式的计划,均已完成了大量此

类型的计算,其结果是直接可用的。在相关部门积累的这些计算结果中,不仅包含了系统的典型工况、重要故障,还隐含了描述系统所必须使用的模型的详细程度等专家知识,其中所包含的信息量对训练 SVM 而言是充分的。文[8]的计算也表明:不用任何额外的数据支持,仅利用这些数据就可以训练出有效的稳定判别分类器。

经计算发现:对于 IEEE39-BUS 系统,经过 CV 算法对核函数的参数进行优化后,不同的核函数所得到的分类器效果相差不多;而对于 IEEE145-BUS 系统,使用高斯核函数的效果最好。有测试表明,在某些系统的计算中,核函数不同则可能对应差异很大的结果^[8]。然而关于核函数如何选择的问题,理论上尚没有统一的方法,即使是针对某一特定问题,大量采用的仍是试探法。这也是 SVM 当前研究的重要问题之一。

信息融合是在不同的若干输入组合上进行的。通过不同组合可以发现哪些组合的效果最好。以下是仿真给出的测试结果。

表 1 仿真结果

Tab. 1 Simulation result

方案序号	分类器融合方案	IEEE39-BUS 系统测试正确率/%	IEEE145-BUS 系统测试正确率/%
1	子分类器 1	86.5	88.0
2	子分类器 2	84.0	82.0
3	子分类器 3	88.0	83.5
4	子分类器 4	84.0	80.5
5	所有输入特征取并集	88.0	85.0
6	子分类器 1、2、3 融合	94.0	93.5
7	子分类器 1、2、4 融合	93.0	91.5
8	子分类器 1、3、4 融合	93.0	93.0
9	子分类器 2、3、4 融合	94.0	92.0
10	子分类器 1、2、3、4 融合	96.0	95.5

由于子分类器的输入特征仅为经过初步分析选择出的,故 4 组子分类器的分类效果均不理想,达不到应用要求的精度。对所有输入特征取并集,其结果并未显著提高分类精度。但经过信息融合后,分类器则被明显强化,分类精度显著提高。

从错误分类结果索引分析中可以发现,融合后的分类器与子分类器的分类结果相比,有如下几种情形:

(1)强校正。融合后的分类器能在少数分类器结果正确而多数分类器结果错误的情形下仍然得到正确结果。以第 6 组分类器融合方案(1、2、3 融合)为例,39-BUS 系统的结果中,子分类器 1、2 均

错分的样本有 10 个,但其中有 6 个在最终融合结果中被校正过来,就是说,子分类器 3 的正确分类校正了子分类器 1、2 的 2 个错误分类结果,少数派获胜,这是融合方法的一大独到优势。在 145-BUS 系统的结果中,该比例为 6/11。强校正现象的原因在于这些样本虽然被子分类器 1、2 错分类,但判别函数的计算表明这些样本均位于分类器 1、2 的类别边界上,在这些区域容易发生错分类;而该样本在分类器 3 中则远离类别边界,在这些区域分类正确的可靠性较高,故最终融合结果仍为正确分类。这一点充分说明了所加入的融合 SVM 具有判断子分类器分类结果有效性的功能。这也是融合能够显著提高分类精度的主要原因。

(2) 反常校正。反常校正是指在所有子分类器均错误分类时,融合后的结果仍然是正确的。造成反常校正现象的根源在于,在训练第 2 层 SVM 的样本中,也存在反常校正的样本,这些样本被分类器“记忆”了下来。在 39-BUS 系统中,上述子分类器 1、2 均错分的另 4 个样本中有 3 个未被校正,仍存在于融合后的错分样本中,还有 1 个被子分类器 3 也错分,但在最终融合结果中却被反常校正为正确分类了。

(3) 过度校正。过度校正是指在只有少数子分类器分类错误而多数子分类器分类正确的情况下,融合后的分类结果却出现错误。过度校正现象也说明子分类器数目并非越多越好。本文中 39-BUS 系统仿真结果的方案 6、8 中及 145-BUS 系统仿真结果的方案 7、8、9 中,各出现了一个过度校正的样本。

(4) 融合失败。融合失败是指在所有子分类器分类结果均正确时,融合后的结果却是错误的。本文仿真中未出现融合失败的情况。

对分类器的互补性能做如下分析:

(1) 被校正的样本数(强校正与反常校正之和)与总的错分样本数的比值。以 39-BUS 系统仿真结果的方案 6 为例,子分类器 1、2 错分被 3 校正过来的比例为 6/10,子分类器 1、3 错分被 2 校正过来的比例为 5/7,2、3 错分被 1 校正的比例为 5/9。可以看出,尽管子分类器 2 的分类效果较差,但它对于校正子分类器 1、3 却能起到相当的作用(5/7,在 145-BUS 系统中的比例为 5/8)。

(2) 融合结果中的错分样本归类。仍以 39-BUS 系统仿真结果的方案 6 为例,最终结果中的 12 个错

分样本中有 5 个是子分类器 1、2、3 均错分的(这 5 组样本有 3 组在子分类器 4 中也被错分并保留在第 10 组融合方案中了,另 2 组被子分类器 4 正确分类并在第 10 组融合方案中校正了过来),另外的 7 个错分样本中的 6 个是两组错分没有被另外一组校正的,另一个是被过度校正而造成错误的。对于其它组合的融合方案,也可以做相同分析。

应指出的是,本文仿真结果中 4 组子分类器一起融合的分类效果最佳,但这并非意味着子分类器数目越多越好。当某些子分类器因输入特征选择不当而造成彼此冲突时,就有可能出现“过度校正”和融合失败数量的增加,造成融合后分类精度的降低。通常寻找最优子分类器数目和组合的方法是测试几种情况后留下最好的,而严格的方法是使用最优脑损伤算法^[19]从一个全连接的网络开始,然后从中去除连接。在网络经过第一次训练之后,用信息论方法确定一个可丢弃的连接的最优选择,网络再次训练,如果性能没有降低,则重复这个过程。在本文提出的 2 层 SVM 融合模型中,最差情形下的计算复杂度为 $O(2^n)$ 。如何寻找计算复杂度更低的有效算法,是一个应进一步研究的课题。

6 结论

为解决支持向量机方法应用于暂态稳定分析中所遇到的输入特征选择难的问题,本文根据信息融合思想构造了 2 层 SVM 组合分类器模型,采取对来自于不同启发知识所构造出的多组输入特征的子分类器计算结果进行信息融合,来降低预测结果对输入特征的敏感性。IEEE 39-BUS 和 IEEE 145-BUS 系统上进行的仿真表明,信息融合方法具有以下明显优势:

(1) 能够以弱规则为基础,经过融合模型的学习得到比原有规则强得多的新规则,经过强化后的规则能够满足工程应用要求;

(2) 即使在少数分类器得到的结果正确而多数分类器结果错误的情形下,融合结果也有可能是正确的,并能保持较高的校正比例,体现了有效域思想。

如何快速地计算出基于哪些组弱规则进行信息融合能最有效地得到强规则,即对哪些子分类器进行信息融合才得到最佳分类器,是一个值得进一步研究的课题。同时,通过分析校正样本,研究各种预测方法的互补性能及各自的有效域,也体现出采用非模型方法分析暂态稳定的优势。

参考文献

- [1] 顾雪平, 曹绍杰, 张文勤. 基于神经网络暂态稳定评估方法的一种新思路[J]. 中国电机工程学报, 2000,20(4): 77-83.
Gu Xueping, Cao Shaojie, Zhang Wenqin. A new framework for transient stability assessment based on neural networks[J]. Proceedings of the CSEE, 2000, 20(4): 77-83.
- [2] 戴仁昶, 张伯明, 戚其芸. 暂态稳定仿真的综合人工智能方法[J]. 中国电机工程学报, 2002,22(12): 1-5.
Dai Renchang, Zhang Boming, Qi Qihui. Transient stability simulation based on artificial intelligence[J]. Proceedings of the CSEE, 2002, 22(12): 1-5.
- [3] 王平洋. 电工数学与智能技术[J]. 电网技术, 1998, 22(5): 1-4.
Wang Pingyang, Electro-mathematics & intelligent technology[J]. Power Technology, 1998, 22(5): 1-4.
- [4] 赵登福, 王蒙, 张讲社, 等. 基于支撑向量机方法的短期负荷预测[J]. 中国电机工程学报, 2002,22(4): 26-30.
Zhao Dengfu, Wang Meng, Zhang Jiangshe *et al.* A support vector machine approach for short term load forecasting[J]. Proceedings of the CSEE, 2002, 22(4): 26-30.
- [5] 董明, 孟源源, 徐长响, 等. 基于支持向量机及油中溶解气体分析的大型电力变压器故障诊断模型研究[J]. 中国电机工程学报, 2003, 23(7): 88-92.
Dong Ming, Meng Yuanyuan, Xu Changxiang *et al.* Fault diagnosis model for power transformer based on support vector machine and dissolved gas analysis[J]. Proceedings of the CSEE, 2003,23(7): 88-92.
- [6] 李元诚, 方廷健, 于尔铿. 短期负荷预测的支持向量机方法研究[J]. 中国电机工程学报, 2003, 23(6): 55-59.
Li Yuancheng, Fang Tingjian, Yu Erkeng. study of support vector machines for short-term load forecasting[J]. Proceedings of the CSEE, 2003,23(6): 55-59.
- [7] 许涛, 贺仁睦, 王鹏, 等. 基于统计学习理论的电力系统暂态稳定评估[J]. 中国电机工程学报, 2003, 23(11): 51-55.
Xu Tao, He Renmu, Wang Peng *et al.* Power system transient stability assessment based on statistic learning theory[J]. Proceedings of the CSEE, 2003, 23(11): 51-55.
- [8] Moulin L S, A P Alves da Silva, El-Sharkawi M A *et al.* Support vector machines for transient stability analysis of large-scale power systems[J]. IEEE Trans. on Power Systems, 2004,19(2): 818-825.
- [9] Muller K R, Mika S, Ratsch G *et al.* An introduction to kernel-based learning algorithm[J]. IEEE Trans. On N.N. 2001, 12(2): 181-201.
- [10] Burges J C. A tutorial on support vector machines for pattern recognition[R]. Bell Laboratories, Lucent Technologies. 1997.
- [11] Vladimir N Vapnik, 张学工译. 统计学习理论的本质[M]. 北京: 清华大学出版社, 2000.
- [12] Efron B. Estimating the error rate of a prediction rule: Improvement on cross-validation[J]. Journal of the American Statistical Association, 1983, 78: 316-331.
- [13] Vladimir N Vapnik, 许建华, 张学工译. 统计学习理论[M]. 北京: 电子工业出版社, 2004.
- [14] 唐巍, 陈学允. BP 算法应用于电力系统暂态稳定分析的新策略[J]. 电力系统自动化, 1997, 21(3): 47-50.
Tang Wei, Chen Xueyun. A new strategy of bp algorithm applying to transient stability analysis of power system[J]. Automation of Electric Power Systems, 1997, 21(3): 47-50.
- [15] 顾雪平, 张文朝. 基于 Tabu 搜索技术的暂态稳定分类神经网络的输入特征选择[J]. 中国电机工程学报, 2002, 22(7): 66-70.
Gu Xueping, Zhang Wenchao. Feature selection by tabu search for neural-network based transient stability classification[J]. Proceedings of the CSEE, 2002, 22(7): 66-70.
- [16] 余贻鑫, 王成山. 电力系统稳定性理论与方法[M].北京: 科学出版社, 1999.
- [17] 冯飞, 余贻鑫. 电力系统功率注入空间的动态安全域[J]. 中国电机工程学报, 1993, 13(3): 14-22.
Feng Fei, Yu Yixin. Dynamic security regions of power system in injection space[J]. Proceedings of the CSEE, 1993, 13(3): 14-22.
- [18] 曾沅, 余贻鑫. 电力系统动态安全域的实用解法[J]. 中国电机工程学报, 2003, 23(5): 24-28.
Zeng Yuan, Yu Yixin. A practical direct method for determining dynamic security regions of electric power systems[J]. Proceedings of the CSEE, 2003, 23(5): 24-28.
- [19] Le Cun Y, Jackel L, Boser B *et al.* Handwritten digit recognition: Applications of neural network chips and automatic learning[J]. IEEE Communications Magazine, 1989,27(11): 41-46.

收稿日期: 2005-02-06。

作者简介:

马 骞 (1978-) 男, 博士研究生, 研究方向为电力系统智能调度;
杨以涵 (1927-) 男, 教授, 博士生导师, 研究方向为电力系统分析与控制;

刘文颖 (1956-) 女, 教授级高工, 研究方向为电力系统智能调度与安全稳定;

齐 郑 (1977-) 男, 博士研究生, 研究方向为电力系统分析与控制;

郭金智 (1979-) 男, 硕士, 从事电力系统调度运行与分析工作。