

电力系统低频振荡模式的自动分类研究

陆超, 陆秋瑜

(清华大学电机系电力系统国家重点实验室, 北京 100084)

摘要: 大规模复杂电网在不同运行条件下其低频振荡模式变化明显, 基于扰动信号和类噪声信号的振荡辨识海量结果需在线分析, 以完成多个模式的自动识别与分类。针对上述问题, 提出并设计了一套不依赖于经验的振荡模式自动分类系统。该系统包括特征选择和分类器两个部分, 利用特征选择实现了大范围的降维, 并对比分析了线性的 Fisher 分类器、非线性的二次型和 k 近邻分类器性能。进一步的, 基于南方电网简化仿真数据进行了验证, 其结果说明了所设计和实现方法的有效性, 为进一步的预警、分群等提供了重要信息。

关键词: 低频振荡; 振荡模式辨识; 模式分类; 特征选择; Fisher 分类; k 近邻分类; 二次型分类

Research on power system low frequency oscillation modes classification

LU Chao, LU Qiu-yu

(State Key Laboratory of Power Systems, Department of Electrical Engineering, Tsinghua University, Beijing 100084, China)

Abstract: In large scale power grids, the low frequency oscillations may change significantly in different operation conditions, and the mass results identified based on ringdown and ambient signals must be analyzed online to complete the automatic classification of multiple oscillation modes. In this paper, a computer-oriented oscillation modes recognition and classification system is designed, which includes feature selection and classifier. The first one is used to select the most sensitive features and reduce the dimensions greatly. About classifier, three linear and nonlinear methods, i.e. Fisher, k-neighborhood and quadratic, are compared. The classification performances are verified based on the China Southern Power Grid simplified data, and the results demonstrate the feasibility and advantages of the proposed approach.

Key words: low-frequency oscillations; oscillation mode identification; pattern classification; feature selection; Fisher classifier; k-nearest neighbor classifier; quadratic classifier

中图分类号: TM712 文献标识码: A 文章编号: 1674-3415(2010)04-0035-04

0 引言

全国电网互联和西电东送工程的推进, 大区电网存在弱交流联系, 而高增益快速励磁系统的普遍采用及传统 PSS 在解决区间低频振荡问题方面的不足, 使得近年来低频振荡事件时有发生, 严重影响了电网的稳定和安全。动态稳定性已成为约束电力系统稳定极限的瓶颈因素之一^[1-2]。

目前广域测量系统可实时监测和记录电网动态过程, 及时发现低频振荡, 这为我们分析和抑制低频振荡提供了基础。研究系统低频振荡首先需要得到主要振荡模式的频率、阻尼比、幅值和初始相位等信息。传统的信号分析方法有傅里叶变换、短

时傅里叶变换、小波分析等, 近年来还出现了 Prony 算法、ARMA 算法等现代谱估计算法。Prony 及其改进算法^[3-4]能够直接计算出电力系统低频振荡信号的振荡特征, 分辨能力高, 主要用于扰动情况下的振荡过程, 已在实际电网中得到了广泛应用; ARMA 算法^[5-6]同样可辨识出系统的主导振荡模式信息, 主要用于处理稳态情况下由负荷随机扰动引起的小幅波动类噪声信号。

目前不同振荡模式的区分主要是在小干扰稳定计算或 Prony 分析基础上利用经验进行, 通过计算可得到系统低频振荡的基本信息, 如振荡的频率、相位角、幅值、阻尼比等, 但由于不同运行方式下不同模式的相位分布或频率分布范围会发生一定变化, 如南方电网云贵模式和云广模式都可能是 0.6 Hz 左右, 而云贵模式和贵州内部模式又都有可能是 0.8 Hz 左右, 这就需要结合多个地点的相位等信息、

基金项目: 电力系统国家重点实验项目 (SKLD08Z01); 中国南方电网有限责任公司重大科技专项

依靠一定的经验加以区分。这一方法可用于离线计算或扰动事后分析等情况，但随着前述可在稳态情况下监视系统动态稳定水平的类噪声信号低频振荡辨识技术的发展，大量低频振荡信息的在线持续分析判断不能再依靠人工进行，需要计算机自动实现不同振荡模式的判断过程。本文针对这一新问题，基于已辨识得到的不同节点、不同振荡模式的频率、阻尼比、幅值和初始相位信息，利用模式识别与分类的方法，实现了不同振荡模式的自动分类处理，为实时监测、进一步的预（告）警和调度决策提供了重要依据，也可以用于简化现有的离线 Prony 辨识结果后期处理。

本文首先介绍了模式分类系统的基本原理和实现方法，包括线性与非线性分类器的设计、特征选择等，实现了对振荡模式的识别与分类的自动化。然后基于南方电网仿真辨识数据，对电网振荡系统中常见的三类问题的模式分类进行了分析，证实了模式分类在电力系统应用的可行性，并比较了不同分类器的适用范围和优缺点，最后提出了适用于电力系统的模式自动分类方法。

1 低频振荡模式自动分类系统设计原理

基于统计方法的模式分类系统主要由四部分组成：数据获取、预处理、特征提取或选择、分类器设计或分类决策^[7]，如图 1。

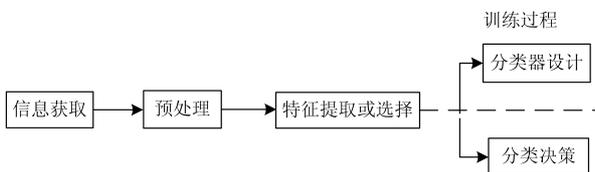


图 1 模式分类系统基本构成

Fig.1 Basic components of pattern classification system

分类系统分两部分：分类器训练和分类决策。对于经过训练的分类器，只需将原始数据输入，即可输出分类决策，即输入样本所属类别。故系统设计的难点在训练过程的特征选取和分类器的设计。

1.1 分类器设计

分类器是模式分类系统中最重要的一环，其思路是在样本训练集的基础上确定某个判决规则，使按这种判决规则对被分类对象进行分类所造成的错误分类率最小或引起的损失最小。Bayes 分类器是在各类别总体的概率分布已知且类别数一定的情况下才可实现的分类器，它以错误分类率为准则函数，故其错误分类率可达到最小，为最优分类器^[7]。

分类器根据其判别函数的形式可分为线性分类器和非线性分类器两种^[7]。

1.1.1 线性分类器

线性分类器是最简单的分类器。它的判别函数为线性函数，分界面为超平面，故容易实现。

线性分类器的判别函数 $g(x)$ 的常用表达式为：

$$g(x) = \omega^T x + \omega_0 \tag{1}$$

其中： x 为 d 维特征向量； ω 为权向量； ω_0 为常数，称为阈值权。

Fisher 分类器是经典的线性分类器模型。寻求变换向量 ω^* 时，它的目标是使得投影后不同类间样本的距离尽可能大，同一类内样本的距离尽可能小，从而形成不同类样本间相互分离，类内各自紧凑的集合。当维数 d 和样本数 N 都足够大时，fisher 分类器为最优分类器。

1.1.2 非线性分类器

非线性分类器的决策面是超曲面，主要应用于复杂界面的区分，常用的有分段线性判别函数和二次判别函数。

分段线性判别函数的决策面是由若干段超平面组成的，故它能逼近各种形状的超曲面，具有很强的适应能力。 k 近邻法是应用较广的分段线性分类法之一，其主要思想是取 x 的 k 个近邻样本，计算这 k 个近邻中多数属于哪一类，就把 x 归为那一类。当 k 趋于无穷时， k 近邻法为最优分类器。 k 近邻法的突出优点是算法简单，错误概率较小，但是它们要求样本的数量大，故计算量和存储量也会很大，在减少计算和存储方面有相当多的改进算法，在此不一一列举。

二次判别函数的适用范围比分段线性函数更广。其判别函数的一般表达式为：

$$g(x) = x^T W x + \omega^T x + \omega_0 \tag{2}$$

其中： W 为 $d \times d$ 实对称矩阵； ω 为 d 维向量。

二次判别函数确定的决策面是一个超二次曲面，包括超球面、超椭球面、超双曲面等，其分界面的复杂程度要大很多。当样本为正态分布，且样本均值向量和协方差矩阵的估计值接近真实的分布值，二次判别函数的错误分类率将逼近 Bayes 分类器的错误率。

1.2 特征选择

在前面讨论分类器时可以看到，在样本数不是很多的情况下，用过多的特征进行分类器的设计，无论从计算的复杂程度还是分类器的性能来看都是不适宜的。所以我们需要在设计分类器之前分析各种特征的有效性，选择可分性较高的特征量进行分类器训练，就将高维特征空间压缩到低维特征

空间, 并获得更好的分类器性能。

对于特征有效性的衡量, 需要一个量化的准则。常用的方法是用不同类别样本间的距离^[8] (称为可分性判据 J) 衡量特征空间的可分性, 距离越大, 两类的分离程度就越大。但是由于特征间的相互联系, 可分性判据较大的前 d 个最有效的特征并非是最优的特征组, 这就涉及到组合的问题。穷举法虽能找到最合适的结果, 但计算量过大。对于特征量的组合情况, 有多种优化的搜索方法, 本文使用的是一种简单的自下而上的搜索方法: 顺序前进法, 即每次从未入选的特征中选择一个特征, 使得它与已入选的特征组合在一起时 J 最大, 直到特征数为 d 。

2 低频振荡模式分类系统的仿真分析

2.1 仿真方法

低频振荡模式的自动分类验证基于南方电网简化仿真数据进行。南方电网的主导区间振荡模式为云贵—广东模式、云南—贵州模式。此外, 在部分运行方式下, 贵州、云南等省网内部一些区内振荡模式也比较明显。

在上述简化的仿真辨识数据中, 在四个区域内选择 20 个母线作为测量节点, 利用现有的 Prony 和 ARMA 等谱估计方法可在暂态和稳态时辨识得到每个节点参与的不同模式的振荡频率、阻尼比、幅值和初始相位信息, 以此作为分类的基础数据。这样, 分类系统的每个样本由一个八十维的特征向量决定其振荡的模式。相关分类计算等采用基于 Matlab 的 PRTOOL 工具箱进行^[9]。

多个不同振荡模式的识别是实际应用时经常会遇到的问题, 考虑到实际电网中的主导振荡模式一般不会超过 3 个, 本文以下部分将以三类问题为例进行说明。

基于本文第 1 节中所述原理, 我们采用的特征选择方法为顺序前进法, 分类器设计采用了线性的 fisher 分类器、分段线性的 k 近邻法和非线性的二次判别法。

选用错误分类率来评价分类器的性能, 取多段时间的样本, 经 10 倍交叉验证^[10]得到该时段样本的错误分类率, 取其平均值作为结果。

2.2 仿真结果及分析

模拟的模式一为云南和贵州之间机组的振荡, 模式二为云南和广东之间的振荡, 模式三为云南内部的振荡。取 20 个样本, 画出每个节点所有样本的原始特征量的分布区间及均值如图 2 (其中, 三个模式分别用实线、点线、点划线表示多个样本的分

布区间, 用圆形、星形、五角星表示样本的均值位置)。

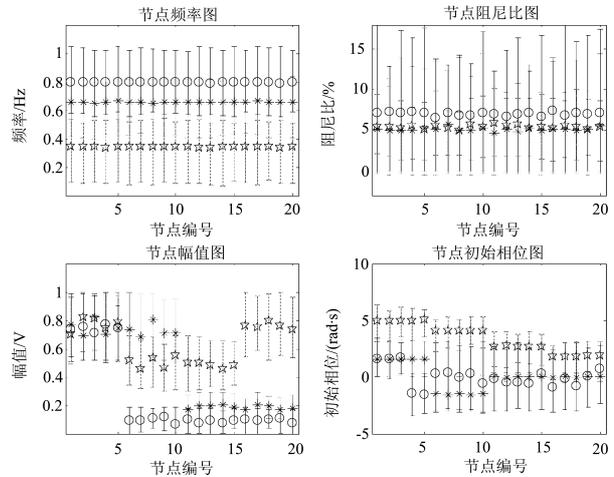


图 2 三类问题原始数据分布图

Fig.2 Raw data distribution graph of three classes

将 20 个节点的特征进行编号, 1~20 为节点振荡频率, 21~40 为阻尼比, 41~60 为幅值, 61~80 为初始相位。通过计算得到:

当选择两维特征时, 筛选得到的特征向量为 57、59, 即第 17、19 个节点的幅值; 当选择五维特征时, 筛选得到的特征向量为 57、59、48、46、62, 即第 17、19、8、6 个节点的幅值和第 2 个节点的初始相位, 这和分布图上观察到的幅值和相位的分布分离性较大的现象相符。

经过两维特征选择后训练的分类器决策面如图 3 (散点图为第 57、59 个特征分布图)。

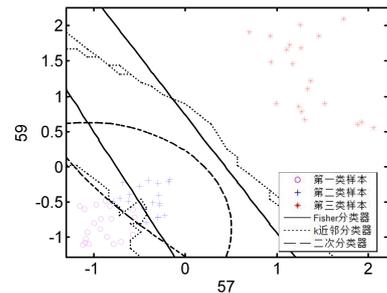


图 3 选择二维特征时分类器的分界面

Fig.3 Boundary of classifiers when choosing two features

从图 3 可以看出, 二次分类器的边界是曲线, k 近邻是分段直线, fisher 是直线。选择两维特征会导致较大的错误率, 将特征维数增至五维, 选择的特征量及由其训练得到的分类器决策面在某两维平面上的投影如图 4。

对比可知, 特征维数为五维时, 样本信息更充分, 在某些特征平面上样本的分离性更大, 分类器

的边界也得到改善，可减小错误分类率。

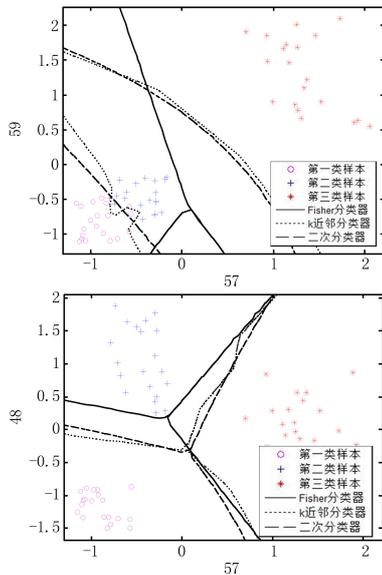


图 4 选择五维特征时分类器的分界面

Fig.4 Boundary of classifiers of when choosing five features

特征选择维数分别为二维、五维、十维时各分类器的错误分类率计算结果如表 1。

表 1 三类问题的各分类器错误概率

Tab.1 Errors for classifiers of three classes

错误概率	Fisher	k 近邻	二次判别
无特征选择	0.083 3	0.160 8	0.058 3
选择二维特征	0.150 0	0.084 2	0.075 8
选择五维特征	0.068 3	0.012 5	0.009 2
选择十维特征	0	0.000 8	0

从结果可以看出，选择合适维数的特征量，既可实现较大范围的降维，也可改善分类器的性能，使错误概率降低到可接受的程度。

当样本数量较少时，非线性分类器比线性分类器更有优势。二次判别法和 k 近邻法对于复杂的边界均表现出了较大的适应能力。由于实际系统中多数样本遵循正态分布，由第 1 节的理论分析可知二次判别的错误分类率能逼近 Bayes 分类器，故无论是高维还是降维，二次分类器的性能都最好。

由前面的分析可知，分类器的错误分类率还与样本的数量有关。将样本数量增至 40，得到各分类器的错误分类率列于表 2。

样本数量增加后，fisher 分类器不需进行特征选择前即可正确地对三类问题进行分类。可以证明当样本特征维数 d 和样本数 N 都很大时，fisher 分类器等价于 Bayes 分类器^[7]，故在样本数量和存储空间充足的情况下，推荐采用 fisher 分类器。另外，

样本数量的增加对 k 近邻分类器和二次判别分类器的性能都有一定程度的改善。

表 2 样本数量为 40 时分类器的错误概率

Tab.2 Errors for classifiers when the number of samples increases to 40

错误概率	Fisher	k 近邻	二次判别
无特征选择	0	0.126 7	0.034 6
选择二维特征	0.137 9	0.080 0	0.067 9
选择五维特征	0.090 4	0.011 7	0.010 4
选择十维特征	0.001 2	0	0

将样本数量增至 60，计算结果对于非线性分类器的改善程度并不明显，故不再增加样本数。

从以上的分析还可以看到，特征维数的选择和分类器的错误分类率并没有直接的对应关系。当维数过少时，信息不够充分，较难代表复杂的边界情况；维数过多，即带来计算和存储的难题，一些无用的噪声特征也会影响分类器的性能。目前关于满意特征选择^[11]的研究，引入了多目标优化的思想，可挑选出具有维数少、复杂度低和分类能力强的满意特征集。特征选择最优化也是我们进一步研究的内容。

3 结论

本文基于电力系统类噪声信号辨识得到的不同主导振荡模式的频率、阻尼比、幅值、相位信息，设计了振荡模式分类系统，包括特征选择及线性、非线性分类器的算法，并结合三类模式的情况，对特征选择及分类算法的性能进行了计算、分析和理论解释，验证了本文方法的正确性和有效性。

在基于稳态类噪声信号的电力系统低频振荡在线辨识与监测中，实测数据和样本充足，维数较大，分类系统建议采用 fisher 分类器。本文设计的低频振荡模式自动分类方法，在完善低频振荡的在线监测和预警功能的同时，也为调度运行人员进一步的潮流调整决策提供了重要基础信息，使其可采取针对性的措施，预防低频振荡事件的发生。

参考文献

[1] 王铁强, 贺仁睦, 王卫国, 等. 电力系统低频振荡机理的研究[J]. 中国电机工程学报, 2002, 22 (2): 21-25. WANG Tie-qiang, HE Ren-mu, WANG Wei-guo, et al. The Mechanism Study of Low Frequency Oscillation in Power System[J]. Proceedings of the CSEE, 2002, 22 (2): 21-25.

(下转第 48 页 continued on page 48)

角的估计测量上, 其精度高的优势更加明显。

3 结论

扩展 Prony 算法以最小二乘法为基础, 采用指数矩阵线性组合可以直接求解信号幅值、相角、阻尼比和频率, 具有较高的实时性和精确度。本文通过理论分析和仿真测试对比, 证明了扩展 Prony 算法应用在直流谐波测量方面是有效的, 并且具有更高的精度。扩展 Prony 算法的这一特点为直流输电谐波分析提供了更好的测量工具。

参考文献

- [1] Arrillaga J, Bradley D A, Bodger P S. Power System Harmonics[M]. New York: John Wiley & Sons, 1985.
 - [2] 金明, 刘远龙. 用于电力系统谐波分析的 ANN 算法[J]. 电网技术, 1997, 21(5): 52-54.
JIN Ming, LIU Yuan-long. A New ANN Harmonic Analysis Algorithm For Power System[J]. Power System Technology, 1997, 21(5): 52-54.
 - [3] Dusan Agrez. Weighted Multipoint Interpolated DFT to Improve Amplitude Estimation of Multi-frequency Signal[J]. IEEE Trans on Instrumentation and Measurement, 2002, 51(2): 287-291.
 - [4] Hildedrand F B. Introduction to Numerical Analysis[M]. New York: McGraw-Hill, 1956.
 - [5] 张贤达. 现代信号处理[M]. 北京: 清华大学出版社, 1995.
ZHANG Xian-da. Modern Signal Processing[M]. Beijing: Tsinghua University Press, 1995.
 - [6] 王峰, 游广增, 黄莹, 等. 高压直流输电 3 脉动谐波电压源计算[J]. 高电压技术, 2009, 35(10): 264-268.
WANG Feng, YOU Guang-zeng, HUANG Ying, et al. Calculation of Three-pulse Harmonic Voltage Source for HVDC Transmission[J]. High Voltage Engineering, 2009, 35(10): 264-268.
 - [7] 王峰, 朱光友, 游广增, 等. 高压直流输电直流侧谐波电流计算[J]. 高电压技术, 2009, 35(9): 229-233.
WANG Feng, ZHU Guang-you, YOU Guang-zeng, et al. Calculation of DC Side Harmonic Currents for HVDC Transmission[J]. High Voltage Engineering, 2009, 35(9): 229-233.
-
- 收稿日期: 2010-01-11; 修回日期: 2010-01-18
作者简介:
王铁松 (1965-), 男, 工程师, 主要从事机电设计研究与电气安全评价工作。E-mail: tiesong85374689@sina.com
-
- (上接第 38 页 continued from page 38)
- [2] Kakimoto Naoto, Ohnogi Yukio, Matsuda Hisao, et al. Low Frequency Oscillation and Field Tests on Power System Stabilizer[A]. in: Electrical Engineering in Japan[C]. 1987. 27-35.
 - [3] Amano M, Watanabe M, Banjo M. Self-testing and Self-tuning of Power System Stabilizers Using Prony Analysis[A]. in: IEEE Engineering Society, Winter Meeting[C]. 1999. 655-660.
 - [4] 肖晋宇, 谢小荣, 胡志祥, 等. 电力系统低频振荡在线辨识的改进 Prony 算法[J]. 清华大学学报: 自然科学版, 2004, 44(7): 883-887.
 - [5] Pierre J W, Trudnowski D J, Donnelly M K. Initial Results in Electromechanical Mode Identification from Ambient Data[J]. IEEE Trans on Power Systems, 1997, 12(3): 1245-1250.
 - [6] Wies R W, Pierre J W, Trudnowski D J. Use of ARMA block Processing for Estimating Stationary Low-frequency Electromechanical Modes of Power Systems[J]. IEEE Trans on Power Systems, 2003, 18(1): 167-173.
 - [7] 边肇祺. 模式识别[M]. 北京: 清华大学出版社, 1988.
 - [8] 王娟, 慈林林, 姚康泽. 特征选择方法综述[J]. 计算机工程与科学, 2005, 27(12): 68-71.
WANG Juan, CI Lin-lin, YAO Kang-ze. A Survey of Feature Selection[J]. Computer Engineering & Science, 2005, 27(12): 68-71.
 - [9] Duin R P W, Juszczak P, Paclik P, et al. A Matlab Toolbox for Pattern Recognition[J]. Delft University of Technology, 2007.
 - [10] Kohavi, Ron. A Study of Cross-validation and Bootstrap for Accuracy Estimation and Model Selection[A]. in: Proceeding of the 14th International Joint Conference on Artificial Intelligence[C]. 1995. 1137-1143.
 - [11] 张葛祥, 金炜东, 胡来招. 满意特征选择及其应用[J]. 控制理论与应用, 2006, 23(1): 19-24.
ZHANG Ge-xiang, JIN Wei-dong, HU Lai-zhao. Satisfactory Feature Selection and Its Applications[J]. Control Theory & Applications, 2006, 23(1): 19-24.
-
- 收稿日期: 2009-09-25
作者简介:
陆超 (1977-), 男, 博士, 副教授, 主要研究方向为电力系统分析与稳定控制。E-mail: luchao@mail.tsinghua.edu.cn