

数据挖掘中适用于分类的时序数据特征提取方法^①

林 珠¹, 邢 延²

¹(广东省计算中心, 广州 510033)

²(广东工业大学 自动化学院, 广州 510006)

摘 要: 特征提取在提高分类的准确性中起着非常关键的作用. 对时序特征提取的方法进行归纳分类, 将有利于对特征提取整体性, 全面性的认识. 回顾现有的时间序列中特征提取的方法, 将其总结为四大类, 它们分别是基于基本统计方法的特征提取、基于模型的特征提取、基于变换的特征提取、基于分形维数的特征提取. 针对每一类的特征提取方法, 进一步研究了它相应的分类方法和它在时间序列数据中的应用邻域.

关键词: 时序数据; 分类; 特征提取

Survey of Feature Extraction Approaches for Time Series Classification

LIN Zhu¹, XING Yan²

¹(Guangdong Computer Center, Guangzhou 510033, China)

²(Guangdong University of Technology, Guangzhou 510006, China)

Abstract: The main contributions of this paper are: 1) The main feature extraction approaches are classified into four categories; 2) The main idea of each category is analyzed, the advantages and disadvantages are pointed out; 3) The guidelines of choosing suitable feature extraction approach is suggested.

Key words: time series; classification; feature extraction

1 引言

时序数据(time series data)广泛存在于现实生活中, 是指同一种现象在不同时间上的相继观察值排列而成的一组数字序列, 其时间轴上的采样值通常又被称为特征^[1]. 时序数据普遍存在于许多重要应用邻域, 比如 DNA 序列、金融数据、传感器网络监控数据、移动对象跟踪数据、机器故障检测数据等等. 由于时序数据与时间相关联, 因而其数据量一般都是非常庞大的, 这就对时序数据挖掘技术提出了更高的要求^[2]. 在时序数据挖掘的研究与应用领域, 时序数据分类是重要任务之一^[1,2]. 例如, 依据语音信号的波形识别出说话人的性别和年龄, 依据心电图的时序波形识别出病者所患的病症, 依据地震波的历史数据, 去识别地震的类型, 依据在机器运转过程中进行故障检测和识别故障类型, 甚至在客户关系管理中根据某段时间的客户

购买信息, 识别不同的消费群体等等.

衡量分类技术优劣的核心指标是分类准确率, 而提高分类准确率途径有两种: 一是改进分类器; 二是采用特征提取技术(feature extraction). 特征提取是在分类前对数据时间采样值上进行适量的归约, 以达到减少数据量同时提高分类准确率(底线是不牺牲分类准确率)的目的.

时间序列除了具有的趋势性、季节性、周期性等一般特征之外, 不同的时序数据又存在不同的个别特征. 如金融数据, 普遍具有“高峰厚尾”和“平方序列微弱而持续的自相关”的特点; 而地震波则具有强度随时序延伸而减弱的特点; 语音信号幅度具有一定的范围, 并以零幅和近零幅的概率高, 而且长时间的语音信号会有相当多的无信号区间, 即所谓的语音寂静区间; 心电信号则具有很强的周期性, 它的主要特征是

^① 基金项目: 广东省科技计划项目基金(2011B060500049, 2010B090400545, 2010A040300006)

收稿时间: 2012-02-06; 收到修改稿时间: 2012-03-04

低电压(0.8~1mV), 小电流(12uA), 重复频率低, 每个波段具有各自的频率. 针对时序数据的这些特征, 所选择的特征提取方法应该能提取出时间序列中具有较好分类能力特征, 进行特征提取后的特征矢量能够很好地代表原有的时间序列数据, 这样才能取得良好的分类效果.

本文对分类中常用的时序特征提取方法归纳和总结, 现阶段特征提取方法主要有四类, 分别是基于基本统计方法的特征提取, 基于模型的特征提取, 基于变换的特征提取以及基于分形理论的特征提取. 针对每一类特征提取方法, 本文总结了各自适用的应用领域, 并提及相应的分类方法.

2 特征提取方法综述

特征提取的方法有很多, 总结起来可以归为四大类: 基于基本统计方法的特征提取, 基于模型的特征提取, 基于变换的特征提取, 基于分形维数的特征提取. 通过这些特征提取后的特征矢量, 能够达到较好的分类效果.

2.1 基于基本统计方法的特征提取

基本统计方法的特征提取, 就是提取数据波形的均值, 方差, 极值, 波段, 功率谱, 过零点等统计特征来代表原有的时序数据作为特征矢量. 时域的常见基本统计特征有均值, 方差, 极值, 过零点, 边界点^[3], 波段的长短峰值等, 而频域的基本统计方法有功率谱, 功率密度比, 中值频率, 平均功率频率等. 在肌电信号(EMG)数据中, 常采用的时序统计特征为过零点, 积分肌电值, 方差等, 而频域中则常采用 EMG 功率谱的平均功率频率和中值频率^[4]. ECG 中则常提取它的 R-R 间隔均值、熵值、变化值、功率谱密度等, 而脑电信号(EEG)则常提取它的峰值, 熵值^[5], 非线性能量等^[6], 或是提取 QRS 波的峰值、波长等, QT 间隔, ST 间隔等统计特征^[6,7].

基于统计特征提取后的特征矢量, 可采用线性判别式构造分类器; 也可采用神经网络进行分类, 此时可达到较好效果^[7,8]. 它适用于信号波形的统计特征比较明显的时间序列数据, 如 EEG 信号、ECG 信号等医学数据.

2.2 基于模型的特征提取

基于模型的特征提取, 是指用模型去刻画时间序列数据, 然后提取模型的系数作为特征矢量去进行分

类. 对于平稳时间序列, 可用通用 ARMA 模型(自回归滑动平均模型)及其特殊情况的自回归模型、滑动平均模型(MA)或组合-ARMA 模型等来进行拟合. 对于非平稳时间序列则要先将观测到的时间序列进行差分运算, 化为平稳时间序列, 再用适当模型去拟合这个差分序列. 当然, 不同的时序数据都会有比较适合它的模型去进行特征提取. 如对于 ECG 数据, 通常采用 AR 模型进行特征提取, 而相对应的 MAR 模型则用于多通道的 ECG 数据中^[9-11]. AR 模型是一个线性的, 二阶矩平稳模型, 比较适合短数据分析, 不仅在 ECG 数据中具有优势, 在肌电信号(EMG)中也是如此. Graupe 在研究用肌电信号进行动作识别时, 证实了 AR 模型比 ARMA 模型更具优势. 近年来的研究也表明, 用 AR 模型进行特征提取时, 识别率是非常之高^[4]. 而对于金融时序数据, 常用的多元线性回归和 ARMA 模型等都不再适合, 针对金融市场的价格波动聚集现象, 应采用 ARCH 模型(自回归条件异方差族计量模型)^[12]; 而采用 SVAR(结构向量自回归)模型更加适合刻画金融数据中常出现的“高峰厚尾”和“弱自相关”现象^[13].

选择合适的模型进行特征提取, 产生的模型系数作为特征矢量. 针对这类特征矢量, 往往采用简单的判别函数即可达到理想的分类效果. 最常用的当属二次判别函数(QDF)^[9,10]. 有时也用到最大似然的判断规则进行分类^[14].

2.3 基于变换的特征提取

通过变换的手段, 使适合分类的特性突显出来, 也是经常用到的特征提取方法. 变换包括时频变换和线性变换. 时频变换中具有代表性的方法是快速傅立叶变换、短时傅立叶变换、倒谱系数等. 而线性变换, 主要有 PCA、ICA、SVD、线性判别式分析、要素分析、映射等等. 其中最具有代表性的是 PCA 和 K-L 变换, 还有著名的小波变换、小波包技术.

2.3.1 基于时频变换的特征提取

时频变换是将信号从时域变换到频域的一种手段. 时间序列数据在特征提取中常采用傅立叶变换、倒谱系数等时频变换方法.

傅里叶变换是将时域的信号变换成频域的信号, 它是把时域的信号变换成由频率, 幅值和相位的正弦波的组合. 将时序数据进行傅立叶变换, 然后选择它的系数作为特征矢量, 若选择前面的系数, 则代表了信号的低频特性, 选择较大的系数, 则代表了信号

的能量特征^[15]。在心电信号 QRS 波识别中用 DFT 有效进行降维和提取特征,然后用神经网络进行分类,其敏感度可高达 98%^[16]。在股市中应用傅立叶变换与反变换,可以有效去除噪声和进行数据约简^[17];亦可通过傅立叶变换求取功率谱,来表征时序数据的统计特征^[18]。

为了减少傅立叶变换的运算量,可运用快速傅立叶变换算法,它利用 DFT 系数的对称性,周期性和可约性等性质将长序列的 DFT 分解为若干个短序列的 DFT 运算,其中,最常用的是 FFT 的表达形式是按时间抽取的基 2 FFT 算法,然后提取频率系数作为特征矢量^[16]。

短时傅立叶变换(Short-Time Fourier Transform, STFT)建立在传统的傅立叶变换基础之上,其基本思想是将信号 $s(t)$ 用窗函数 $w(t)$ 截断,并让该窗函数沿着信号滑动,于是分为一个个短时信号,并对这些短时信号分别进行快速傅立叶变换进行频谱分析,从而能更精确地观察出信号的频谱变化情况。窗函数的长短选择决定了时频特征体现的强弱(窗较宽时频域性较强,时域性较弱)^[19]。

特征提取还经常用到倒谱系数也作为信号的特征矢量。倒谱分析又称为二次频谱分析,它分为实倒频谱和复倒频谱。其中实倒谱又为功率倒频谱,它是先将时序信号进行离散傅立叶变换,然后取自然对数,再实行离散傅立叶反变换。复倒谱是将时序信号通过 Z 变换以后取对数,再求反 Z 变换而得到的。倒谱系数应用最广泛的领域是在语音信号中^[20-23],包含梅尔倒谱和线性预测倒谱,其识别效果优于其它特征提取方法,其中梅尔倒谱又要优于线性倒谱^[24]。针对倒谱提取出来的特征矢量,语音信号常用 HMM 模型进行分类识别。

2.3.2 基于线性变换的特征提取

线性变换中有很多特征提取方法颇为经典,特别是 PCA 和小波变换。

PCA 为主成分分析,它与独立成分分析 ICA, K-L 变换,奇异值分解极为相似,有殊途同归的效果,在数据降维中都经常用到。研究表明,利用 PCA 变换可以在信息损失最小的前提下,用较少的分量代替原来的高维数据,达到降维的效果,从而使得处理数据的时间和费用大大降低。另一方面,由于各主成分是相互垂直的,所以增大了类间距,减小了类内差

异,可提高分类精度^[25,26]。但 PCA 不能解决非线性特征提取,对于非线性数据,提出了核主成分分析的观点。它是通过一个非线性映射,把输入数据映射到一个高维的特征空间,在特征空间上进行线性主成分分析^[27,28],其优点是要求的数据量小,并具有很好的抗噪能力^[29]。

应用 PCA 或核主成分分析进行数据降维后,往往会进一步进行特征提取,使数据进一步压缩^[30,31]。特别是核主成分分析,特征提取后的维数仍会很高,应进一步找寻降维的方法。用 PCA 进行特征提取后的特征矢量,常采用最近邻方法^[32,33]、HMM 模型进行分类^[33],特别是采用神经网络^[34,35],可达到较高的分类准确率。即使用最简单的 BP 网络,也有较满意的结果^[36]。

小波变换是将信号分解到不同尺度的线性变换。基于小波变换的特征提取有多种,如提取模极大值特征,能量特征,熵特征,以及适应性小波网络的特征提取等^[37]。如在 ECG 信号的特征提取中,最常用的是基于模极大值的特征提取它蕴含了原始信号的大部分重要信息,可以将其尺度参数 S,平移参数 T 及其幅值作为心电信号的特征量^[38]。

小波变换在所有的时序数据中都得到广泛应用,但对于不同的时序数据进行特征提取时,应选择不同的小波基,常用的小波基有: Haar 小波基, Daubechies 小波基系列, Coiflets 小波基系列, Symlets 小波基系列, Biorthogonal 小波基系列等等。通常我们选择小波函数时应遵守两条原则: (1)对于复数信号应用复数母小波; (2)母小波的形状应与被分析的信号相类似。在对心电信号进行模极大值特征提取时,通常选择 Mexican hat 小波函数^[39],这是因为它的形状与被分析信号相类似以及较好的时域分辨率。而对于股市数据,用具有良好的正交性和紧支撑性的 Daubechies 小波系、Coiflets 小波系、Symlets 小波系会比其他小波系效果好^[40]。而地震波中, Dubieties 小波族中的 Db4 是很适合进行岩石声信号处理^[41];小波包技术不同于小波技术,它将频率轴分成不同大小的区间,这一特点特别适合地震波,因为地震波强度随着传递频率明显衰减。在地震波特征提取中运用小波包技术优于傅立叶变换^[42]。同时,可以将小波系数和核函数方法相结合进行特征提取,识别率高达 100%^[43]。

对于小波技术提取出的特征矢量,可以采用平面

鉴别法, 距离分类器等, 但最常见的分类方法是神经网络和 SVM, 并且都能取得较好的分类效果^[44-46].

2.4 基于分形理论的特征提取

分形是指具有: 无限精细、非常不规则、无穷自相似结构和非整数的点集. 在大自然中, 海岸线、雪花、云雾这些不规则形体都属于分形, 即部分与整体有自相似性^[47]. 分形学与混沌学息息相关, 成为非线性科学的两大重要组成部分, 分形理论真正发展只有十余年, 但应用于特征提取已越来越广泛, 特别是在时间序列中. 但分形理论进行特征提取时, 并不是所有信号都适合. 要看信号在某个尺度下是否具有可分形特征, 即不同状态下的分形维数是不同的, 这样才具有可分性. 如故障诊断中, 将信号分为 N 个状态, 每个状态可提取一个分形维数作为特征参数, 把这个分维数与学习好的故障分类结果比较, 可判断是否有故障发生^[48].

分形理论用于特征提取时, 主要是针对非线性信号, 是用它的定量分析指标分维数作为特征矢量. 分维数是指空间的扩展程度, 可以用分数来表示, 它的提出可定量很多有争议的数据, 比如一直测不准的英国海岸线^[25]. 分维数有多种, 其中相似维数, 适用于严格相似的一小类集合豪斯道夫维数难于计算, 盒维数在稠密可列集上与集所在空间维数相等^[47], 对于一维的时间序列, 关系维是最简单实用的方法^[48]. 分形理论是研究非线性时序变量的有力武器, 可揭示出复杂系统下的早、洪灾害时序变化的内在规律^[49], 也可用于股票离群点的检测^[50], 采用多重分形特征对金融管理或风险预测数据进行分类^[51], 分维数用于心博分类时, 准确率可高达 99.49%^[52].

分形维数只得到一个数值作为特征矢量, 利用此特征矢量, 用贪婪算法, 或设定阈值即可进行分类, 当然, 它作为一个非线性参数, 也可以跟李雅谱诺指数, 小波技术提到的参数一起组成特征矢量进行分类^[53].

2.5 小结

上述四大类的特征提取方法在分类过程中都得到了广泛的应用. 但对于具体的应用领域, 应选择较为适合该领域数据的特征提取方法, 才能较好地提取该数据的特征. 而对于不同的特征提取方法进行提取后的特征矢量, 也有较为适合并常用的分类方法. 因此, 对于上述方法, 可以总结为表 1 所示:

表 1 各类特征提取方法

类别	特征矢量	主要方法	分类方法	常用邻域
基本统计方法	基本统计特征作为特征矢量	均值, 方差, 极值, 波段, 功率谱, 过零点等	线性判别式分类器、神经网络分类器	统计特征较明显的时间序列, 如 EEG、ECG 等医学数据
基于模型	模型的系数作为特征矢量	自回归滑动平均模型、滑动平均模型、组合-ARMA 模型、自回归条件异方差计量模型等。	利用判别函数构造分类器, 常用二次判别函数。	医学数据、金融数据、地震波等自然现象数据
基于变换	时频	利用变换后的频域系数等作为特征矢量 傅立叶变换、倒谱系数	经傅立叶变换后分类方法广泛; 倒谱系数则常采用 HMM 模型进行分类。	医学数据、语音信号数据
	线性	主成分或小波系数等作为特征矢量 PCA、K-L 变换、小波变换	PCA; 最近邻、神经网络等; 小波变换; 神经网络和 SVM	广泛, 小波变换更适合地震波分类
基于分形理论	分形维数作为特征矢量	相似维数、盒维数、关系维等	贪婪算法, 或设定阈值即可	非线性信号

3 结论与展望

综上所述, 这四大类的特征提取方法都广泛应用于时序数据, 而且各有千秋, 对于具体数据, 可以选择其中比较适合它的特征提取, 如股票时序, 往往采用模型的方法会比较能表征它的波动特征. 基于统计特征的方法最早得到发展, 也最为简单, 但在信号有强噪声的情况不是很适合, 而基于变换的方法, 无论是小波变换还是主成分分析, 在当今应用得最为广泛, 可以达到很好的分类效果. 而针对信号常常出现的非线性情况, 我们可以采用分形理论去解决, 提取信号的分形维数进行分类.

考虑到时序信号的复杂性, 可以采用多种特征提取方法的结合, 如求倒谱系数时可以结合到 LPC 线性

预测方法,而通常我们也会把小波技术跟分形理论结合。多种方法的综合将会是未来研究特征提取的主要方向。

参考文献

- 1 陈晓云,吴本昌,韩海涛.基于多维时间序列数据挖掘的降雨天气模型研究.计算机工程与设计,2010,4.
- 2 Mark AKHB. Last Data Mining In Time Series Databases, World Scientific,2004.
- 3 韦东兴,陈晓云,徐荣聪.基于角点检测的图像形状特征提取方法.计算机工程,2010,4.
- 4 加玉涛,罗志增.肌电信号特征提取方法综述.电子器件,2007:326-330.
- 5 刘慧,谢洪波,卫星.基于模糊熵的脑电睡眠分期特征提取与分类.数据采集与处理,2010,25(4):484-489.
- 6 Greene BR, Boylan GB, Reilly RB. Connolly Combination of EEG and ECG for improved automatic neonatal seizure detection, Clin Neurophysiol 2007,118:1348-1359.
- 7 Gao GQ. Computerised detection and classification of five cardiac conditions,AUT University.
- 8 商卫波.心电信号自动分析与诊断处理方法研究.西安:西北工业大学,2005.
- 9 葛丁飞,邵宇权,蒋惠忠.基于双导联 Ecg 和多变量回归模型的远程心电诊断算法研究.航天医学与医学工程,2004:355-359.
- 10 G.D.-F.H.B.-P.X. Xin-Jian. Study of feature extraction based on autoregressive modeling in ECG automatic diagnosis. Acta Automatica Sinica, 2007,33(5).
- 11 葛丁飞,夏顺仁.Ar 模型在远程心电诊断中的应用.中国生物医学工程学报,2004:222-229.
- 12 钱争鸣.Arch 族计量模型在金融市场研究中的应用.厦门大学学报(哲学社会科学版),2000:126-129.
- 13 余素红,张世英.Sv 与 Garch 模型对金融时间序列刻画能力的比较研究.系统工程,2002:28-33.
- 14 Kalantzis T, Papanastassiou D. Classification of GARCH time series: an empirical investigation. Applied Financial Economics, 2008,18:759-764.
- 15 Mrchen F. Time series feature extraction for data mining using DWT and DFT, 2003.
- 16 Valenza G, Lanata A, Ferro M, Scilingo EP. Real-time discrimination of multiple cardiac arrhythmias for wearable systems based on neural networks. Computers in Cardiology, 2008.1053-1056.
- 17 郭躬德,王晖.DavidBell 时间序列数据分析与预处理.小型微型计算机系统,2003:2228-2232.
- 18 江近仁,孙景江.以反应谱和功率谱密度函数表征的强震运动的统计特性.中国地震,1994:327-340.
- 19 Amini AM. Qualitative features extraction from sensor data using Short-Time Fourier Transform, Visual Information Processing XIV 5817(2005)289-295.
- 20 Holmberg M, Gelbart D, Hemmert W. Automatic speech recognition with an adaptation model motivated by auditory processing, Audio, Speech, and Language Processing. IEEE Trans. on, 2006,14:43-49.
- 21 Xu H, Tan ZH, Dalsgaard P, Lindberg B. Exploitation of spectral variance to improve robustness in speech recognition. Electronics Letters, 2006,42:312-314.
- 22 F.H.a.H. Ney Noise Level Normalization and Reference Adaptation for Robust Speech Recognition. International Workshop on Automatic Speech Recognition: Challenges for the New Millennium. 2000.
- 23 Hilger F, Ney H. Quantile based histogram equalization for noise robust large vocabulary speech recognition, Audio, Speech, and Language Processing. IEEE Trans. on, 2006,14:845-854.
- 24 高瑞华.多种预处理方法在语音检测中应用效果的比较研究.控制理论与控制工程[硕士学位论文].浙江工业大学,2004.64.
- 25 J.R.G. Townshend,C.O.Justice and V. Kalb Characterization and Classification of South American Land Cover Types Using Satellite Data. International Journal of Remote Sensing, 1987,8:1189-1207.
- 26 邢玉娟,谭萍,李明.一种新的说话人识别序列特征提取方法.兰州理工大学学报,2009,35(4):98-102.
- 27 Scholkopf SAB, Muller KR. Nonlinear component analysis as a kernel eigenvalue problem. Neural Computaion.
- 28 杨绍华.一种基于核主成分分析的人脸识别方法.河北科技师范学院学报,2008:45-48,62.
- 29 周小程,马向玲,范洪达,庞文强.基于核函数分类的多维时序特征选择方法应用.光电与控制,2010,17(7):74-77.
- 30 侯振雨,蔡文生,邵学广.主成分分析-支持向量回归建模方法及应用研究.分析化学,2006:617-620.

- 31 楼蒋. Method of Feature Extraction Suitable for Hyperdimensional Time Series Data. 浙江科技学院学报, 2007, 19(2).
- 32 胡利平, 刘宏伟, 吴顺君. 基于两级 2dpca 的 Sar 目标特征提取与识别. 电子与信息学报, 2008: 1722-1726.
- 33 李志农, 曾明如, 韩捷, 何永勇. 主分量分析和因子隐 Markov 模型在机械故障诊断中的应用. 机械强度, 2007: 25-29.
- 34 朱帮助. 基于特征提取的选择性神经网络集成方法. 计算机科学, 2008: 132-133, 172.
- 35 刘爱霞, 刘正军, 王静. 基于 Pca 变换和神经网络分类方法的中国森林制图研究. 长江流域资源与环境, 2006: 19-24.
- 36 王海燕, 刘鲁, 刘玲. 基于 Gra 和 pca 的 Bp 神经网络应用研究. 管理评论, 2007: 50-54.
- 37 张静远, 张冰. 基于小波变换的特征提取方法分析. 信号处理, 2000: 156-162, 155.
- 38 马陈, 马国亮. 心电图的小波变换特征提取. 中国科技信息, 2006(1A).
- 39 Andreao RV, Dorizzi B, Boudy J. ECG signal analysis through hidden Markov models. IEEE Trans. on Biomedical Engineering, 2006, 53: 1541-1549.
- 40 胡博, 卫宏儒, 廖福成. 小波变换在股票分析中的应用. 系统管理学报, 2007: 365-369.
- 41 金解放, 赵奎, 王晓军, 赵康. 岩石声发射信号处理小波基选择的研究. 矿业研究与开发, 2007: 12-15.
- 42 朱江梅, 吴兴方. 小波包变换压缩与地震数据处理. 中国海上油气(工程), 2006: 169-173.
- 43 田喜英. 主元分析和非线性方法在表面肌电信号中的应用研究[硕士学位论文]. 上海交通大学, 2009.
- 44 Gupta AMR, Singh K. A Time-Series-Based Feature Extraction Approach for Prediction of Protein Structural Class. EURASIP. Journal on Bioinformatics and Systems Biology, 2008.
- 45 Palancz B, Benyo B. Classification of time series via wavelet subband analysis using support vector machine classifier. Periodica Polytechnica Electrical Engineering 50, 2006: 129-140.
- 46 郑继明, 邢峰, 吴渝, 李婧. 基于小波变换和支持向量机的音频分类. 重庆邮电大学学报(自然科学版), 2008: 212-216.
- 47 王春. 基于小波和分形理论的齿轮故障特征提取及噪声的和谐化研究. 机械工程, 重庆大学, 2006, 4(1): 1-145.
- 48 王庆华, 张兴彪, 张洪朋, 孙玉清. 分形理论在液压泵故障诊断中的应用. 大连海事大学学报自然科学版, 2004: 40-43.
- 49 蔡爱民, 查良松. 基于分形理论的安徽省旱、洪涝灾害时序特征分析. 安徽农业大学学报, 2005: 546-550.
- 50 孙金花, 冯英俊, 胡健. 基于分形理论的股票时序数据离群模式挖掘研究. 运筹与管理, 2008: 135-140.
- 51 关腾, 许娜. 金融时间序列的多重分形分类. 郑州大学学报(理学版), 2008, 4.
- 52 Raghav S, Mishra AK. Fractal feature based ECG arrhythmia classification. IEEE, Piscataway, NJ, USA, 2008. 5.
- 53 钟维年, 高清维, 陈燕玲. 基于小波和多重分形的金融时间序列聚类. 系统工程, 2009: 58-61.