



(12)发明专利申请

(10)申请公布号 CN 109034260 A  
(43)申请公布日 2018.12.18

(21)申请号 201810889309.9

(22)申请日 2018.08.07

(71)申请人 东南大学

地址 211102 江苏省南京市江宁区东南大学路2号

(72)发明人 杨路宽 孙立

(74)专利代理机构 南京苏高专利商标事务所  
(普通合伙) 32204

代理人 柏尚春

(51) Int. Cl.

G06K 9/62(2006.01)

B01D 53/50(2006.01)

B01D 53/80(2006.01)

F04D 27/00(2006.01)

权利要求书2页 说明书7页 附图3页

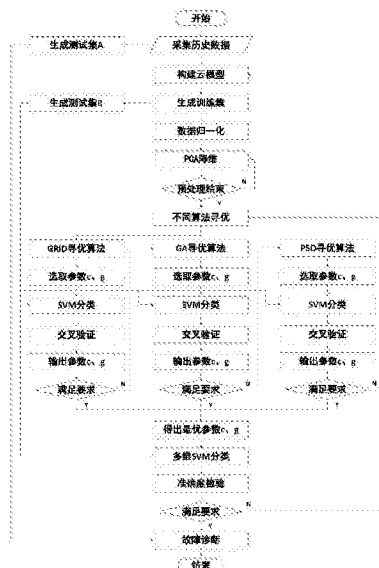
(54)发明名称

基于统计原理和智能寻优的脱硫塔氧化风机故障诊断系统及方法

(57)摘要

本发明公开了一种基于统计原理和智能寻优的脱硫塔氧化风机故障诊断系统及方法,属于故障诊断领域。系统包括数据采集单元、数据预处理单元、参数优化单元、支持向量机分类单元、故障诊断单元和人机交互单元;诊断方法步骤如下:利用云模型对历史运行数据进行扩充生成训练集、将训练集数据归一化并降维、采用网格化寻优、遗传理论寻优、粒子群寻优算法完成对支持向量机重要参数的设定、采用多维支持向量机对降维的数据分类、检验分类准确率并反馈到云模型构建过程和算法寻优过程中直至准确率达到预设要求、实现对实时数据的故障诊断。本发明实现了脱硫塔氧化风机是否处于故障状态的诊断,诊断准确度高、自学习性强,具有实际应用价值。

CN 109034260 A



1. 一种基于统计原理和智能寻优的脱硫塔氧化风机故障诊断系统,其特征在于:包括:  
数据采集单元,所述的数据采集单元以脱硫系统和主机系统的DCS控制系统所记录的历史数据作为故障诊断系统的历史数据来源,通过支持OPC协议的标准OPC方式建立历史数据库,并通过DCS控制系统所采集的实时数据不断扩充历史数据库;

数据预处理单元,所述的数据预处理单元通过生成正向云反应发生器、数据归一化和主元分析降维将历史运行数据处理成适合分析的已预处理数据并分别保存于不同数据集;

参数优化单元,所述的参数优化单元通过网格化寻优算法、遗传理论寻优算法和粒子群寻优算法处理并寻找多维支持向量机最优的惩罚参数和核函数参数;

支持向量机分类单元,所述的支持向量机分类单元利用向量机分类原理将数据预处理单元生成的数据进行分类和检验;

故障诊断单元,所述的故障诊断单元以脱硫塔进出口二氧化硫浓度、循环浆液泵流量、各台浆液循环泵电机电流为参考数据,调用支持向量机分类单元已经训练好的相关参数,对系统氧化风机是否存在故障进行判断;

人机交互单元,所述的人机交互单元利用MATLAB GUI工具箱创建用户可视化图形界面,显示设定时间段内各台浆液循环泵电机电流、各台浆液循环泵喷浆量、烟道出口SO<sub>2</sub>浓度的变化趋势,并显示故障诊断的结果。

2. 根据权利要求1所述的基于统计原理和智能寻优的脱硫塔氧化风机故障诊断系统,其特征在于:所述数据预处理单元采用云模型生成新数据集,对历史运行数据中的异常工况和非正常运行工况的数据进行扩充,为后续各单元提供全工况数据集,参数优化单元利用网格化寻优算法、遗传理论寻优算法和粒子群寻优算法内部不同的寻优机制依次对多维支持向量机最优的惩罚参数和核函数参数进行寻优识别。

3. 一种采用如权利要求1所述的基于统计原理和智能寻优的脱硫塔氧化风机故障诊断系统的诊断方法,其特征在于:包括以下步骤:

步骤(1),数据预处理单元接收一个或多个物理量随着时间变化所产生的一组或多组离散数据;

步骤(2),数据预处理单元根据数据使用先后的不同将所接收的一组或多组离散数据按预先设置的比例随机抽取后生成测试集A并保存,所述的测试集A作为故障诊断单元中整体测试集的一部分参与故障诊断单元数据输入过程;将所采集的一组或多组离散历史数据作为输入引入云模型并通过正向云模型发生装置产生足够的和测试集A具有相同统计学特征的离散数据作为数据预处理单元的训练集;

步骤(3),所述的训练集按预先设置的比例随机抽取后生成测试集B并保存,所述的测试集B作为参数优化单元中的测试集引入SVM分类中参与惩罚函数参数c和径向基核函数参数g的寻优过程;

步骤(4),所述的训练集归一化处理并采用PCA主元分析法降维,分析识别已归一化数据的相互关联程度从而识别出对分类效果影响不大的相关物理量并按照预先设置的前n个主成分总占比数对分类效果影响较大的物理量进行主成分占比计算;

步骤(5),预先设置的前n个主成分总占比计算值满足设定值时数据预处理单元结束并进入参数优化单元,预先设置的前n个主成分总占比计算值不满足设定值时对归一化后数据重新进行主成分降维计算直至预先设置的前n个主成分总占比计算值满足设定值;

步骤(6),参数优化单元分别采用网格化寻优算法、遗传理论寻优算法、粒子群寻优算法对SVM的两个重要参数惩罚函数参数和核函数参数寻优从而实现较高的SVM方法分类准确率;

步骤(7),参数优化单元得到的最佳参数 $c$ 、 $g$ 作为SVM方法参数设置的重要部分用于指导多维SVM分类,分类准确性通过引入多维SVM数据预处理单元所生成的测试集B来检验;检验符合要求则支持向量机分类单元结束进入实时故障诊断单元,检验符合不要求则返回至数据预处理单元的构建云模型阶段通过正向云发生器重新产生新云滴数据后按照上述步骤依次重新计算直至支持向量机单元检验符合要求;

步骤(8),经过训练得到分类准确率较高的支持向量机相关参数适用于故障诊断的场合,采用作为运行真实数据的测试集A加以仿真。

4.根据权利要求3所述的采用基于统计原理和智能寻优的脱硫塔氧化风机故障诊断系统的诊断方法,其特征在于:步骤(1)中,所述的数据预处理单元实时监测前脱硫系统运行过程中氧化风机的故障问题,保证气液比始终在合理范围内,确保氧化风机长时间处于零故障运行。

5.根据权利要求3所述的采用基于统计原理和智能寻优的脱硫塔氧化风机故障诊断系统的诊断方法,其特征在于:步骤(2)中,所述的云模型通过计算数据集数学期望、数据集熵、数据集熵的熵这三类指标扩充原有数据集,增加数据集丰富性。

6.根据权利要求3所述的采用基于统计原理和智能寻优的脱硫塔氧化风机故障诊断系统的诊断方法,其特征在于:步骤(4)中,所述PCA主元分析法降维方法对各台浆液循环泵电机电流、各台浆液循环泵喷浆量、烟道出口 $SO_2$ 浓度能反应出脱硫系统脱硫效率的物理量和氧化风机振动程度进行关联度分析,采用PCA降维方法减少未处理数据集维度从而使处理后的数据集适合于参数优化单元的分析。

7.根据权利要求3所述的采用基于统计原理和智能寻优的脱硫塔氧化风机故障诊断系统的诊断方法,其特征在于:步骤(6)中,所述的网格化寻优算法的迭代算法均采用龙格-库塔算法。

8.根据权利要求3所述的采用基于统计原理和智能寻优的脱硫塔氧化风机故障诊断系统的诊断方法,其特征在于:步骤(7)中,所述的SVM方法利用核函数将出于低维度的数据集投射到高维空间中,并在高维空间中利用SVM方法对其分类,以解决数据集在低维空间线性不可分的问题。

## 基于统计原理和智能寻优的脱硫塔氧化风机故障诊断系统及方法

### 技术领域

[0001] 本发明涉及数据分析、设备故障分析、设备故障诊断领域,尤其涉及基于统计原理和智能寻优的脱硫塔氧化风机故障诊断系统及方法。

### 背景技术

[0002] 燃煤发电厂所用的原煤中含有硫元素,原煤在锅炉高温环境下燃烧时,其中的硫元素会转化成含有硫元素的气体从而跟随烟气排入大气,造成对周围大气和水质环境的污染。火力发电厂运行过程中的烟气脱硫是防控环境污染的重点,锅炉运行时会产生大量烟气,烟气中含有煤中硫分在锅炉中经过化学反应所产生的二氧化硫和硫化氢等硫化物,含硫烟气会对环境尤其是大气造成巨大的污染,燃煤发电厂中的烟气容易在下雨的过程中造成酸雨,对水质的污染也很严重。因此如何做到烟气脱硫的技术成本和经济成本的有机统一是目前电力行业普遍需要面对和解决的问题。

[0003] 烟气中的二氧化硫被浆液吸收,与石灰石反应生成不稳定亚硫酸钙或亚硫酸氢钙,为了将亚盐氧化成硫酸盐,需要氧气的参与,氧气就来源于氧化风机鼓入的空气。氧化风机大部分采用罗茨型风机,通过两个或者三个转子,互相挤压,将空气吸入并挤出,原理类似于齿轮泵。氧化风机是否正常运行影响了脱硫系统中二氧化硫和石灰石的平均反应时间和反应充分程度,进而影响机组脱硫效率。

[0004] 石灰石-石膏湿法脱硫系统复杂设备繁多,现场运行环境恶劣,故障出现时仅依靠现场运行人员经验很难准确判断故障原因并找出故障点。为了解决这个问题,现场一般采用一套历史记录系统随时记录系统的参数随时间的变化,在有问题时,技术人员可以调用历史数据帮助故障分析。在电厂脱硫系统中,常规的基于数据的故障诊断方法是基于历史运行数据通过智能算法和聚类方法实现对历史数据的分类,从而对实时数据信号中找出故障特征提供借鉴。但是常规脱硫方法只基于历史运行数据,对于历史运行中没有产生的数据的故障诊断准确度并不高。

### 发明内容

[0005] 发明目的:针对当前脱硫系统运行过程中氧化风机X、Y向振动分析和诊断问题,以及保证脱硫过程中氧气供给充足这一要求,本发明提出一种基于统计原理和智能寻优的脱硫塔氧化风机故障诊断系统及方法。

[0006] 技术方案:本发明基于统计原理和智能寻优的脱硫塔氧化风机故障诊断系统包括:

[0007] 数据采集单元,以脱硫系统和主机系统的DCS控制系统所记录的历史数据作为故障诊断系统的历史数据来源,通过支持OPC协议的标准OPC方式建立历史数据库,并通过DCS控制系统所采集的实时数据不断扩充历史数据库;

[0008] 数据预处理单元,通过生成正向云反应发生器、数据归一化和主元分析降维等手

段将历史运行数据处理成适合分析的已预处理数据并分别保存于不同数据集；

[0009] 参数优化单元,通过网格化寻优算法、遗传理论寻优算法和粒子群寻优算法处理并寻找多维支持向量机最优的惩罚参数和核函数参数；

[0010] 支持向量机分类单元,利用向量机分类原理将数据预处理单元生成的数据进行分类和检验；

[0011] 故障诊断单元,故障诊断单元以脱硫塔进出口二氧化硫浓度、循环浆液泵流量、各台浆液循环泵电机电流实时数据作为输入,调用支持向量机分类单元已经训练好的相关参数,对系统氧化风机是否存在故障进行判断；

[0012] 人机交互单元,利用MATLAB GUI工具箱创建用户可视化图形界面,显示设定时间段内各台浆液循环泵电机电流、各台浆液循环泵喷浆量、烟道出口二氧化硫浓度等物理量的变化趋势,并显示故障诊断的结果。

[0013] 数据预处理单元采用云模型生成新数据集的方法,对历史运行数据中较为稀缺的异常工况和非正常运行工况的数据进行了扩充,为后续各单元提供全工况数据集,参数优化单元利用网格化寻优算法、遗传理论寻优算法和粒子群寻优算法内部不同的寻优机制依次对多维支持向量机最优的惩罚参数和核函数参数进行寻优识别,与只依赖单个寻优算法相比具有寻优效果好、寻优精度高的特点。

[0014] 采用基于统计原理和智能寻优的脱硫塔氧化风机故障诊断系统的诊断方法,包括以下步骤：

[0015] 步骤1,以脱硫系统和主机系统的DCS控制系统所记录的历史数据作为故障诊断系统的历史数据来源,通过支持OPC协议的标准OPC方式建立历史数据库,并通过DCS控制系统所采集的实时数据不断扩充历史数据库。数据预处理单元接收一个或多个物理量随着时间变化所产生的一组或多组离散数据；

[0016] 步骤2,数据预处理单元根据数据使用先后的不同将所接收的一组或多组离散数据按预先设置的比例随机抽取后生成测试集A并保存,测试集A作为故障诊断单元中整体测试集的一部分参与故障诊断单元数据输入过程。将所采集的一组或多组离散历史数据作为输入引入云模型并通过正向云模型发生装置产生足够的和测试集A具有相同统计学特征的离散数据作为数据预处理单元的训练集。

[0017] 步骤3,训练集按预先设置的比例随机抽取后生成测试集B并保存,测试集B作为参数优化单元中的测试集引入支持向量机分类中参与惩罚函数参数 $c$ 和径向基核函数参数 $g$ 的寻优过程。

[0018] 步骤4,训练集归一化处理并采用PCA主元分析法降维,分析识别已归一化数据的相互关联程度从而识别出对分类效果影响不大的相关物理量并按照预先设置的前 $n$ 个主成分总占比数对分类效果影响较大的物理量进行主成分占比计算。

[0019] 步骤5,预先设置的前 $n$ 个主成分总占比计算值满足设定值时数据预处理单元结束并进入参数优化单元,预先设置的前 $n$ 个主成分总占比计算值不满足设定值时对归一化后数据重新进行主成分降维计算直至预先设置的前 $n$ 个主成分总占比计算值满足设定值。

[0020] 步骤6,参数优化单元分别采用网格化寻优算法、遗传理论寻优算法、粒子群寻优算法对支持向量机的两个重要参数惩罚函数参数和核函数参数寻优从而实现较高的支持向量机方法分类准确率。

[0021] 步骤7,参数优化单元得到的最佳参数c、g作为支持向量机方法参数设置的重要部分用于指导多维SVM分类,分类准确性通过引入多维支持向量机数据预处理单元所生成的测试集B来检验。检验符合要求则支持向量机分类单元结束进入实时故障诊断单元,检验不符合要求则返回至数据预处理单元的构建云模型阶段通过正向云发生器重新产生新云滴数据后按照上述步骤依次重新计算直至支持向量机单元检验符合要求。

[0022] 步骤8,经过训练得到分类准确率较高的支持向量机相关参数适用于故障诊断的场合,采用作为运行真实数据的测试集A加以仿真,结果表明经过数据预处理单元、参数优化单元、支持向量机分类单元后多维支持向量机故障诊断准确度高。

[0023] 步骤1中,数据预处理单元实时监测前脱硫系统运行过程中氧化风机的故障问题,保证气液比始终在合理范围内,确保氧化风机长时间处于零故障运行。

[0024] 步骤2中,云模型通过计算数据集数学期望、数据集熵、数据集熵的熵这三类指标扩充原有数据集,增加数据集丰富性。

[0025] 步骤4中,PCA主元分析法降维方法对各台浆液循环泵电机电流、各台浆液循环泵喷浆量、烟道出口SO<sub>2</sub>浓度等能反应出脱硫系统脱硫效率的物理量和氧化风机振动程度进行关联度分析,采用PCA降维方法减少未处理数据集维度从而使处理后的数据集适合于参数优化单元的分析。

[0026] 步骤6中,网格化寻优算法,其迭代算法均采用龙格-库塔算法。

[0027] 步骤7中,SVM方法利用核函数将出于低维度的数据集投射到高维空间中,并在高维空间中利用SVM方法对其分类,可解决数据集在低维空间线性不可分的问题。

[0028] 有益效果:与现有技术相比,本发明中云模型的使用弥补了历史运行数据点不完备的缺点,可以提供更为丰富的离散数据,为后续的过程提供数据数量的支撑;多种寻优方法的使用提供了支持向量机最佳的一些参数,最后结合云模型的数据集和多种寻优方法的最佳参数采用支持向量机理论实现对实时数据的故障诊断。云模型使用提高了本发明故障诊断的适用性,网格化寻优算法、遗传寻优算法、粒子群寻优算法的使用提高了支持向量机故障诊断的准确性,可以为故障诊断提供一套准确可靠的方法。

## 附图说明

[0029] 图1是本发明故障诊断系统示意图;

[0030] 图2是本发明故障诊断方法的流程示意图;

[0031] 图3是基于本发明诊断方法设计出的粒子群寻优算法模型示意图。

## 具体实施方式

[0032] 如图1所示,该系统包括数据采集单元、数据预处理单元、参数优化单元、支持向量机分类单元、故障诊断单元和人机交互单元。支持向量机分类单元得到的c、g通过闭环反馈回路1引入参数优化单元,通过闭环反馈回路2控制故障诊断单元的精度。

[0033] 如图2所示,数据采集单元以脱硫系统和主机系统的DCS控制系统所记录的历史数据作为故障诊断系统的历史数据来源,通过支持OPC协议的标准OPC方式建立历史数据库,并通过DCS控制系统所采集的实时数据不断扩充历史数据库。数据预处理单元通过生成正向云反应发生器、数据归一化和主元分析降维等手段将历史运行数据处理成适合分析的已

预处理数据并分别保存于不同数据集；参数优化单元通过网格化寻优算法、遗传理论寻优算法和粒子群寻优算法处理并寻找多维支持向量机最优的惩罚参数和核函数参数；支持向量机分类单元利用向量机分类原理将数据预处理单元生成的数据进行分类和检验；故障诊断单元以脱硫塔进出口二氧化硫浓度、循环浆液泵流量等参数为参考数据，调用支持向量机分类单元已经训练好的相关参数，对系统氧化风机是否存在故障进行判断；人机交互单元利用MATLAB GUI工具箱创建用户可视化图形界面，显示设定时间段内各台浆液循环泵电机电流、各台浆液循环泵喷浆量、烟道出口SO<sub>2</sub>浓度的变化趋势，并显示故障诊断的结果，增加本方法可视化程度和使用便捷程度。

[0034] 采用基于统计原理和智能寻优的脱硫塔氧化风机故障诊断系统的方法，包括以下步骤：

[0035] 步骤1，以脱硫系统和主机系统的DCS控制系统所记录的锅炉负荷、循环浆液泵流量、脱硫塔进出口二氧化硫浓度等20个物理量作为故障诊断系统的历史数据来源，通过支持OPC协议的标准OPC方式建立历史数据库，并通过DCS控制系统所采集的实时数据不断扩充历史数据库。数据预处理单元接收一个或多个物理量随着时间变化所产生的一组或多组离散数据，以锅炉负荷、循环浆液泵流量、脱硫塔进出口二氧化硫浓度等20个物理量2000组数据作为初始数据输入数据预处理单元；

[0036] 步骤2，数据预处理单元根据数据使用先后的不同将所接收的一组或多组离散数据按预先设置的比例随机抽取后生成测试集A并保存，测试集A作为故障诊断单元中整体测试集的一部分参与故障诊断单元数据输入过程。将所采集的一组或多组离散历史数据作为输入引入云模型并通过正向云模型发生装置产生足够的和测试集A具有相同统计学特征的离散数据作为数据预处理单元的训练集。

[0037] 云模型数学模型如下：正向云发生器的触发机制为首先生成以 $E_n$ 为期望，以 $H_e$ 为方差的正态随机数 $E_n'$ ；生成以 $E_x$ 为期望，以 $E_n'^2$ 为方差的正态随机数 $x$ ；采用钟型函数计算确定度函数， $(x, \mu)$ 是相对于论域 $U$ 的一个云滴；重复上述步骤直至生成足够的云滴。各个数学符号计算公式为（其中 $X$ 为原始数据的样本均值、 $E_x$ 为原始数据的数学期望、 $E_n$ 为云模型的熵、 $S$ 为原始数据的方差、 $H_e$ 为云模型熵的熵）：

$$[0038] \quad E_x = \bar{X}$$

$$[0039] \quad E_n = \sqrt{\frac{\pi}{2}} \times \frac{1}{n} \sum_{i=1}^n |x_i - E_x|$$

$$[0040] \quad H_e = \sqrt{S^2 - E_n^2}$$

[0041] 云模型的数学模型建立后，将锅炉负荷、循环浆液泵流量、脱硫塔进出口二氧化硫浓度等20个物理量2000组数据作为云模型中原始数据即原始的云滴，运用正向云反应发生器对20个物理量分别生成10000组云滴作为训练集。

[0042] 步骤3，训练集按预先设置的比例随机抽取后生成测试集B并保存，测试集B作为参数优化单元中的测试集引入支持向量机分类中参与惩罚函数参数 $c$ 和径向基核函数参数 $g$ 的寻优过程。

[0043] 步骤4，训练集归一化处理并采用主成分分析法降维，分析识别已归一化数据的相互关联程度从而识别出对分类效果影响不大的相关物理量并按照预先设置的前 $n$ 个主成分

总占比数对分类效果影响较大的物理量进行主成分占比计算。

[0044] 主成分分析法数学模型如下:对原始数据进行标准化处理;计算样本相关系数矩阵;计算相关技术矩阵R的特征值和相应的特征向量;选择重要的主成分,并写出主成分的表达式;计算主成分权重;依据主成分所占的权重得出经过主成分分析法处理后的数据。

[0045] 步骤5,预先设置的前n个主成分总占比计算值满足设定值时数据预处理单元结束并进入参数优化单元,预先设置的前n个主成分总占比计算值不满足设定值时对归一化后数据重新进行主成分降维计算直至预先设置的前n个主成分总占比计算值满足设定值。

[0046] 步骤6,参数优化单元分别采用网格化寻优算法、遗传理论寻优算法、粒子群寻优算法对支持向量机的两个重要参数惩罚函数参数和核函数参数寻优从而实现较高的支持向量机方法分类准确率。利用MATLAB GUI工具箱创建用户可视化图形界面,显示设定时间段内各台浆液循环泵电机电流、各台浆液循环泵喷浆量、烟道出口二氧化硫浓度等物理量的变化趋势,并显示氧化风机的故障分析和诊断结果。

[0047] 网格化寻优算法结合数据预处理单元输出的二维离散数据,将惩罚参数和核函数参数均预先设置成 $2^{-8}$ 到 $2^8$ 的范围,采用定步长搜索方法找出惩罚参数和核函数参数最优值,若网格化寻优算法得出的惩罚参数和核函数参数最优值在设定误差范围内,则将其作为支持向量机分类单元参数,若不满足则转入遗传理论寻优算法阶段。

[0048] 对遗传理论寻优算法结合数据预处理单元输出的二维离散数据,将惩罚参数和核函数参数依次进行编码、解码、交配、突变、倒位、个体适应度评估、复制七大步骤,寻找惩罚参数和核函数参数最优值,若遗传理论算法得出的惩罚参数和核函数参数最优值在设定误差范围内,则将其作为支持向量机分类单元参数,若不满足则转入粒子群寻优算法阶段。

[0049] 如图3所示,粒子群寻优算法结合数据预处理单元输出的二维离散数据,将惩罚参数和核函数参数作为粒子群的初始化粒子,设置惯性因子、加速常数、最大迭代次数和算法终止误差;评价每一组离散数据的初始适应值;将初始适应值看成每个粒子的局部最优值并将各适应值所对应的位置作为每个粒子初始时局部最优值的所在的位置,将其视作初始的全局最优值;根据公式

$$[0050] \quad v_i^d = w_i v_i^d + c_1 r_1 (p_i^d - x_i^d) + c_2 r_2 (p_g^d - x_i^d) \quad i=1,2,3,\dots$$

[0051] 其中 $v_i^d$ 是d时刻粒子的速度, $w_i$ 是惯性权重, $x_i^d$ 是d时刻粒子的位置, $p_i$ 是个体极值, $p_g$ 是全局极值, $r_1$ 和 $r_2$ 是介于(0,1)之间的随机数。 $c_1$ 、 $c_2$ 是惩罚函数参数,也称为学习因子。

[0052] 更新每个粒子当前的移动速度并对每个粒子的移动速度进行限幅处理,使之不能超过设定的最大移动速度;根据公式 $x_i^d = x_i^d + \alpha v_i^d$ 更新每个粒子所在的位置并将此时的局部最优值作为当前时刻全局最优值;重复上述步骤直到所计算得出的全局最优值满足最大迭代次数;输出所对应的惩罚参数和核函数常数若粒子群寻优算法得出的惩罚参数和核函数参数最优值在设定误差范围内,则将其作为支持向量机分类单元参数,若不满足则转入数据预处理阶段重新构建云模型并重复步骤2到步骤6直到得出的惩罚参数和核函数参数最优值在设定误差范围内。

[0053] 步骤7,参数优化单元得到的最佳参数c、g作为支持向量机方法参数设置的重要部分用于指导多维SVM分类,分类准确性通过引入多维支持向量机数据预处理单元所生成的测试集B来检验。检验符合要求则支持向量机分类单元结束进入实时故障诊断单元,检验符



合不要求则返回至数据预处理单元的构建云模型阶段通过正向云发生器重新产生新云滴数据并重复步骤2到步骤6直到支持向量机单元分类效果符合要求。

[0054] 支持向量机分类原理如下:为了寻找数据的最优分类超平面,采用支持向量机对数据预处理单元产生的离散数据进行分类,通过确定分类函数、最大化分类间隔、确定目标函数、转化目标函数、对转化后的目标函数进行二次规划、引入拉格朗日函数、转化为对偶问题、采用SMO算法求解对偶问题八个步骤完成找到最优分类超平面并完成对原始数据的分类。

[0055] 多维支持向量机的数学模型如下:

[0056] (1) 建立目标函数  $f(x) = w \cdot x + b$ , 其中  $w \cdot x$  为两个一维矩阵的内积,  $b$  为常数;

[0057] (2) 结合目标函数和约束条件将二分问题转化为  $\max \frac{1}{\|w\|}, s.t., y_i(w^T x_i + b) \geq 1, i = 1, \dots, n$ ,

其中  $\|w\|$  为  $w$  的内积,  $y_i$  为目标函数;

[0058] (3) 将上述问题转化成  $\max \frac{1}{2} \|w\|^2, s.t., y_i(w^T x_i + b) \geq 1, i = 1, \dots, n$ ;

[0059] (4) 引入拉格朗日函数  $L(w, b, a) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^n a_i (y_i (w^T x_i + b) - 1)$  和函数  $\theta(w) = \max L$

$(w, b, a)$ , 将目标函数变成  $\min_{w, b} \theta(w) = \min_{w, b} \max_{a_i \geq 0} L(w, b, a) = p^*$ , 其中  $a_i$  为拉格朗日系数,  $p^*$  为引入拉格朗日函数后新的目标函数;

[0060] (5) 利用对偶原理转化为  $\max_{a_i \geq 0} \min_{w, b} L(w, b, a) = d^*$ ; 分别对  $w, b$  求偏导数并令其等于 0, 其中  $d^*$  为利用对偶原理转化后新的目标函数:

$$[0061] \quad \frac{\partial L}{\partial w} = 0 \Rightarrow w = \sum_{i=1}^n \alpha_i y_i x_i$$

$$[0062] \quad \frac{\partial L}{\partial b} = 0 \Rightarrow w = \sum_{i=1}^n \alpha_i y_i = 0$$

[0063] 以上结果代入上述的  $L$ :

$$[0064] \quad L(w, b, a) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^n \alpha_i (y_i (w^T x_i + b) - 1)$$

[0065] 得到:

$$[0066] \quad L(w, b, a) = \frac{1}{2} \sum_{i, j=1}^n \alpha_i \alpha_j y_i y_j x_i^T x_j - \sum_{i, j=1}^n \alpha_i \alpha_j y_i y_j x_i^T x_j - b \sum_{i=1}^n \alpha_i y_i + \sum_{i=1}^n \alpha_i$$

$$[0067] \quad L(w, b, a) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i, j=1}^n \alpha_i \alpha_j y_i y_j x_i^T x_j$$

[0068] (6) 添加KKT条件, 即最优化问题所必须满足的求解条件后求对  $a$  的极大值, 即求解关于所述的对偶问题最优化问题:

$$[0069] \quad \max \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i, j=1}^n \alpha_i \alpha_j y_i y_j x_i^T x_j, s.t., \alpha_i \geq 0, i = 1, \dots, n \text{ 且 } \sum_{i, j=1}^n \alpha_i y_i = 0$$

[0070] (7) 引入核函数和SMO算法对所述的对偶优化问题求解;

[0071] (8) 采用一对多分类原理注意分类从而求解多维分类问题；

[0072] (9) 经过训练得到分类准确率较高的支持向量机相关参数适用于故障诊断的场合,采用作为运行真实数据的测试集A加以仿真,结果表明经过数据预处理单元、参数优化单元、支持向量机分类单元后多维支持向量机故障诊断准确度高。



图1

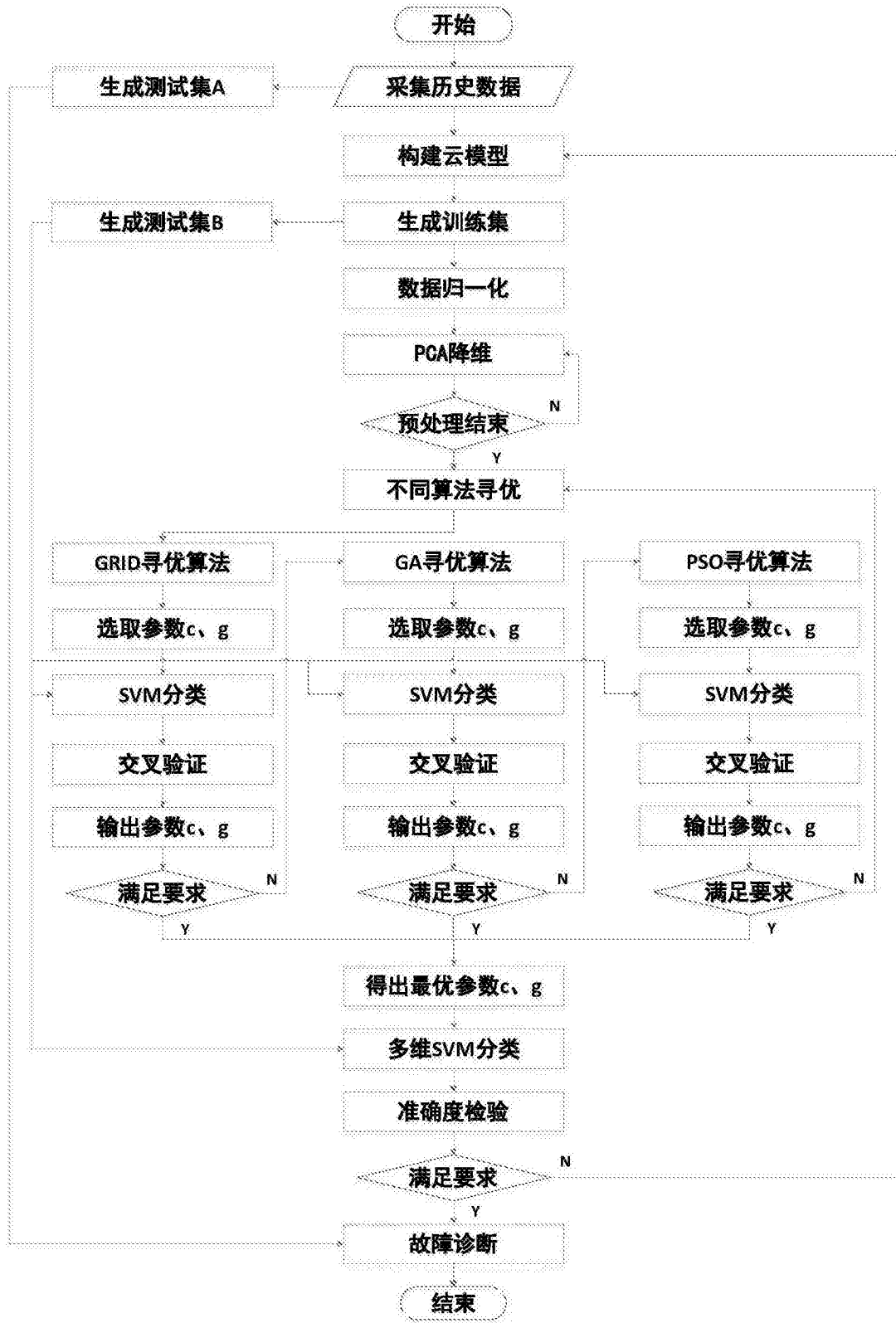


图2

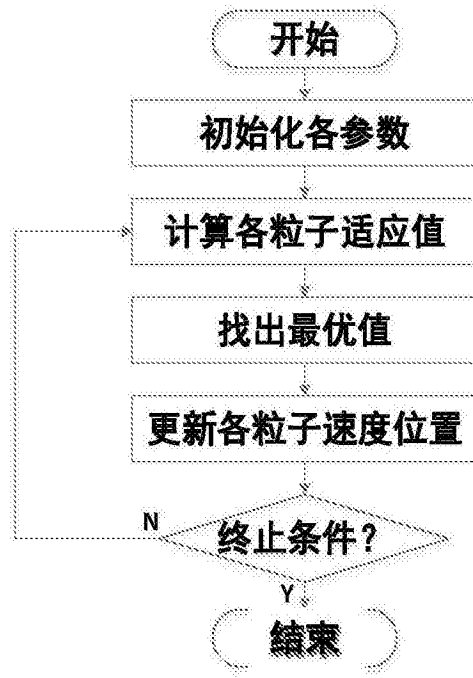


图3