



(12)发明专利

(10)授权公告号 CN 106295692 B

(45)授权公告日 2019.07.12

(21)申请号 201610638135.X

(22)申请日 2016.08.05

(65)同一申请的已公布的文献号

申请公布号 CN 106295692 A

(43)申请公布日 2017.01.04

(73)专利权人 北京航空航天大学

地址 100191 北京市海淀区学院路37号

(72)发明人 何益海 何珍珍 谷长超 韩笑

(74)专利代理机构 北京慧泉知识产权代理有限公司 11232

代理人 王顺荣 唐爱华

(51)Int.Cl.

G06K 9/62(2006.01)

G06Q 50/04(2012.01)

(56)对比文件

CN 104462846 A, 2015.03.25,

CN 104820716 A, 2015.08.05,

张沛朋等.基于PCA-SVM的滚动轴承故障诊断研究.《组合机床与自动化加工技术》.2015,(第11期),第88-90页.

Chih-Wei Hsu等.A Comparison of Methods for Multiclass Support Vector Machines.《IEEE TRANSACTIONS ON NEURAL NETWORKS》.2002,第13卷(第2期),第415-425页.

Zhenzhen He等.Root causes identification approach based on association rule mining for product infant failure.《2015 8th International Symposium on Computational Intelligence and Design》.2015,第624-628页.

M.Demetgul.Fault diagnosis on production systems with support vector machine and decision trees algorithms.《International Journal of Advanced Manufacturing》.2012,(第67期),第2183-2194.

审查员 赵志洪

权利要求书3页 说明书9页 附图2页

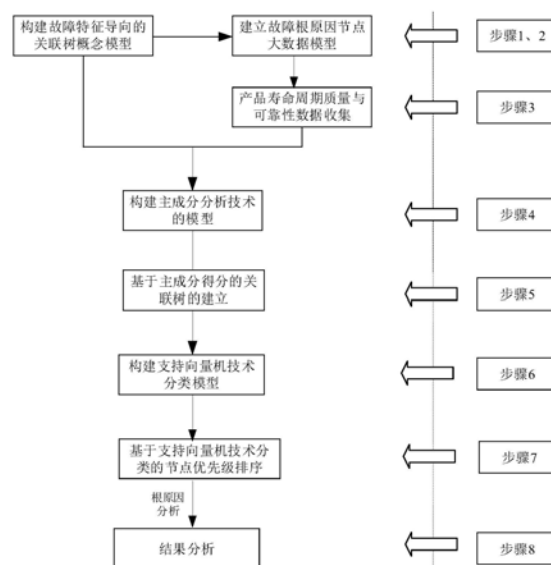
(54)发明名称

基于降维与支持向量机的产品早期故障根原因识别方法

(57)摘要

一种基于降维与支持向量机的产品早期故障根原因识别方法,其步骤如下:1、构建故障特征导向的关联树概念模型;2、构建故障根原因节点大数据模型;3、产品寿命周期质量与可靠性数据收集;4、构建主成分分析技术模型;5、基于主成分得分的关联树构建;6、构建支持向量机技术分类模型;7、基于支持向量机技术分类的节点优先级排序;8、结果分析。本发明从产品可靠性形成角度大数据出发,突破了高维大数据冗余难题,从根本上弥补传统意义上忽略高维数据特征空间而导致的对早期故障机理认识的不精确与误判,提高了大数据环境下早期故障根原因识别效率和准确性,在工程应用中为制造商提供了明确的目标和对象去实施积极的早期故障管控策略。

略。



1. 一种基于降维与支持向量机的产品早期故障根原因识别方法,提出的基本假设如下:

假设1可靠性设计方案不变;

假设2产品设计、制造到使用环节过程是可测量的;

假设3过程检测值相互独立;

假设4销售、运输环节对产品早期故障的影响不考虑;

基于上述假设,其特征在于:该识别方法的实施步骤如下:

步骤一 构建故障特征导向的关联树概念模型;其构建的做法如下:

借鉴公理化域映射理论与瀑布式分解理论,从系统论角度自上而下将早期故障特征映射到功能域,实现功能分解;进而,每个功能需求转化到物理域完成物理分解,继而,被映射到过程域,完成工艺分解并最终形成故障特征导向的关联树概念模型;

步骤二 构建故障根原因节点大数据模型;所述的“故障根原因”是关联树中的关键节点,是指区别于直接原因且具有潜在性的特点,是导致故障发生的最根本的原因;所述的“构建故障根原因节点大数据模型”,其构建的方法如下:

首先,针对目前对早期故障机理的笼统归因于设计、制造和使用无法明确定位的问题,立足于产品早期故障症状,从产品最终可靠性形成角度出发,从设计、制造和使用进行系统梳理导致产品早期故障的原因,提取相关数据;最后,依据所提取的数据信息,建立起潜在故障根原因大数据模型;

步骤三 产品寿命周期质量与可靠性数据收集;是利用历史数据和专家经验手段基于步骤2的大数据模型收集设计、制造、使用三方面的产品生命周期内的质量检测历史数据以及故障数据;

步骤四 构建主成分分析技术模型;其构建方法如下:

首先,在关联树概念模型中选择进行数据降维的目标节点;其次,针对建立的故障根原因大数据模型,将原始数据集的多个维度进行整理,作为训练样本;

给定一个d维空间的具有n个过程变量的N个数据样本, $X=[x_1, x_2, \dots, x_N]$ 是原始故障数据输入向量集, $\bar{x}$ 是样本均值,其协方差计算为:

$$R = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(x_i - \bar{x})^T$$

给定一个n维特征向量 x_i ,其中, $i=1, 2, \dots, N$,特征向量VR的正交投影方向为:

$$RV = \Lambda V$$

其中 Λ 是特征值矩阵,特征值能通过解下面的方程获得:

$$|R - \lambda I| = 0$$

其中 λ 是特征值,其中, $i=1, 2, \dots, N$,I是单位矩阵:

$$\Lambda = \text{diag}[\lambda_1, \lambda_2, \dots, \lambda_k]$$

继而,通过变换把原始的数据通过该变换,变换到一个新的空间中,在这个新的变换空间,能够找到原始数据的主要的属性,通过对这些属性进行选取前d个分量降到d维空间,降低数据的维度;

步骤五 基于主成分得分的关联树构建;其构建方法如下:

把原始数据进行数据的标准化处理,各PCA提取主成分,得到各主成分下的新数据,即为主成分所占的分值,再利用主成分得分值确定各个节点建立故障关联树,其中主成分得分的具体形式如下:

$$\begin{pmatrix} F_{11} & F_{12} & \cdots & F_{1k} \\ F_{21} & F_{22} & \cdots & F_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ F_{n1} & F_{n2} & \cdots & F_{nk} \end{pmatrix};$$

步骤六 构建支持向量机技术分类模型;其构建方法如下:

支持向量机技术是基于结构风险最小化理论之上在特征空间中建构最优分割超平面;假设输入向量集为 $T = (x_i, y_i)$,构造最优超平面的问题转化为下列二次规划问题:

$$\min_{w,b} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \xi_i$$

$$\text{St. } y_i (w^T x_i + b) \geq 1 - \xi_i, \xi_i \geq 0$$

这里 w 和 b 为因子, ξ_i 是松弛变量, C 为惩罚因子;该最优化问题转化为新的二元问题为:

$$\max_{\alpha} \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N y_i y_j \alpha_i \alpha_j x_i^T x_j$$

$$\sum_{i=1}^N \alpha_i y_i = 0, \alpha_i \geq 0$$

在特征空间,为了使最大化边界 $2/\|w\|$ 找到最优超平面,决策函数转换为:

$$f(x) = \text{sgn} \left\{ \sum_{i=1}^N \alpha_i y_i K(x, x_i) + b \right\}$$

这里 x 表示识别的样本, b 为阈值, α_i 为拉格朗日乘数, $\text{sgn}(u)$ 为符号函数, $K(x_i, x_j)$ 为核函数,选择高斯径向基函数为核函数,表达如下:

$$K(x, x_i) = \exp(-\gamma \|x - x_i\|^2)$$

γ 为高斯核函数宽度,其中, $\gamma > 0$,最终问题能转换为如下方程:

$$\max_{\alpha_i} \left\{ \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j K(x_i, x_j) \right\}$$

利用支持向量机处理多分类数据,因此一对多的分类方法需要建立 $K(K-1)/2$ 个分类器;

步骤七 基于支持向量机技术分类的节点优先级排序;这是针对步骤5中估算的关联树节点主成分构建的故障关联树,利用步骤6构建的支持向量机模型进行故障关联树节点的分类;首先选择故障关联树中的目标节点;再根据支持向量机模型给定不同参数 C 和 γ 计算支持向量机分类器的准确率;最后,依据分类的准确率大小得到节点优先级排序;

步骤八 结果分析;是指依据步骤7的结果分析对比不同参数下分类器的结果,得到故障关联树节点权重优先级,并最终确定产品早期故障的根本原因;

通过以上步骤,实现了产品寿命周期高维大数据环境下的故障根本原因的溯源分析,在工程应用中为制造商实施积极的早期故障管控策略提供了明确的目标,能避免对早期故障

原因的误判分析和控制的无效,有利于产品在设计、工艺设计阶段采取事前预防和整改措施。

2.根据权利要求1所述的基于降维与支持向量机的产品早期故障根原因识别方法,其特征在于:该识别方法的使用方法如下:

步骤1 构建故障特征导向的关联树概念模型;利用公理化域映射与瀑布式分解理论构建早期故障特征导向的关联树概念模型,在此过程中不考虑运输和使用的情况下,构建从设计到制造的功能域、物理域、工艺域的故障关联树概念模型;

步骤2 构建故障根原因节点大数据模型;依据早期故障内涵,从产品可靠性最终形成出发,建立起以早期故障为目的,制造过程为核心的,上游追溯到设计、下游立足于使用的大数据模型;

步骤3 从产品最终可靠性形成角度出发,收集产品寿命周期质量与可靠性数据;

步骤4 基于构建的故障关联树概念模型,从针对一个具体故障症状如何从海量数据中找到主节点构建起故障关联树出发,考虑到降维的必要性,建立起主成分分析技术模型;

步骤5 根据主成分分析技术模型,估算节点的主成分得分,即为故障关联树各层级的节点,完成故障特征导向的关联树的构建;

步骤6 针对故障关联树的节点,构建支持向量机技术分类模型辅助分析;

步骤7 基于建立的支持向量机技术模型,利用一对多分类器的准确率得到节点相对效率值,并对得到的节点优先级排序;

步骤8 讨论分析节点权重,针对关联树节点的分类准确率得到产品早期故障的根原因。

基于降维与支持向量机的产品早期故障根原因识别方法

技术领域

[0001] 本发明提供一种基于降维与支持向量机的产品早期故障根原因识别方法,它涉及一种基于降维与支持向量机的产品早期故障根原因的分析方法,属于可靠性建模与分析技术领域。

背景技术

[0002] 过高的早期故障一直是制约装备批产的技术瓶颈,产品经设计到制造末端输出的被制造产品进入使用早期,受环境应力等作用下表现出高的早期故障率,受到顾客的广泛关注。如何从制造和设计的源头对早期故障进行优化和快速消峰没有很好的办法,其关键就在于缺乏从可靠性形成角度系统开展产品早期故障机理的研究,而早期故障机理研究的复杂性表现在影响早期故障的因素多涉及到海量数据,传统意义上快速高效的从高维大数据中提取有效信息缺乏应有的关注,导致对故障根原因的定位不精确。

[0003] 早期故障根原因识别涉及多个环节,通过产品寿命周期的视角,能看到早期故障机理与设计、制造和使用的偏差累积相关。故障机理涵盖的数据包括产品规格配置、质检数据、维修数据和其他相关设计参数和制造参数。因此,从设计、制造和使用中获得的数据能用于识别故障根原因,进而提高产品可靠性。数据的多样性反应了大数据特性,数据规模大,数据种类多,数据要求处理速度快,数据价值密度低的特性。在产品寿命周期存在着大量与故障机理相关的数据,数据单元存储量级达到了TB级或PB级。而且,数据类型包括数据表、文档、音频、图片等。在产品的不同阶段,由于数据量大的特性,导致即使处理大量数据会出现困难。为了有效的识别产品早期故障根原因,我们需要有效的提出大数据信息。然而目前对于产品早期故障的认识比较分散,传统对早期故障机理认识仅笼统归因于设计、制造和使用差错,没有从数据角度深入挖掘故障根原因。本发明聚焦于根本原因分析中的根本原因识别,毫无疑问产品投入使用后,早期故障阶段的产品可靠性问题最为突出,受到研究的关注。由于对早期故障的早期具体定位存在争议导致传统对早期故障阶段的认识比较分散,现有研究已经认识到产品在使用后逐渐暴露出由设计因素、制造因素、材料因素、环境因素等系列问题引发的缺陷,呈现较高的故障率,并具有迅速下降的趋势特征,然而这一认识没有从产品设计到制造再到使用环节层层递进的系统关联因素。鉴于产品可靠性源于设计,成于制造,体现在使用阶段。因而,从系统层面上分析早期故障的影响因素的基础上,考虑制造过程质量偏差在内的从产品生产制造全过程的各环节影响因素的关联分析入手,识别相关关键参数,为突破早期故障机理已有认识尤其重要。

[0004] 由于制造过程的复杂性存在众多不受控制的操作因素,被制造产品通常表现出格外高的早期故障率,产品早期故障的根原因的识别已成为制造商的挑战性问题。尤其在大数据时代,产品寿命周期的大量数据更容易获取,那些高维大数据总是带着很多不相关的噪声信息,导致不仅准确性不明显,而且利用当前的小数据驱动的方法使得模型训练时间出现冗余。此外,传统的面向分析技术的小数据不适用于新的大数据环境。因此,本发明提供了一种基于降维与支持向量机的产品早期故障根原因识别方法,首先考虑传统对早期故

障机理的笼统认识,借鉴公理化域映射理论与瀑布式分解理论构造出一个早期故障特征指引下的分级故障关联树模型指导的潜在根原因的分析。其次,考虑到故障根原因分析所涉及的产品周期大数据的高维数据冗余特征带来的建树不精确问题,利用主成分分析对寿命周期高维数据进行降维处理,以便于得到的关键变量作为潜在的根原因候选集被提取出来辅助故障关联树节点的确定与构建。最后,利用支持向量机技术展开对各潜在故障根原因进行分类识别,以高维数据的视角实现量化分析早期故障根原因,确定出不同层级间影响早期故障发生的关键根原因。

发明内容

[0005] (1) 本发明的目的:

[0006] 过高的早期故障一直是制约装备批产的技术瓶颈,产品经设计到制造末端输出的被制造产品进入使用早期,受环境应力等作用下表现出高的早期故障率受到顾客的广泛关注。当前对早期故障的认识都比较分散,大数据背景下如何将纷繁杂乱的信息集成系统的分析产品早期故障,进而定位早期故障根原因,为研制生产阶段开展积极主动的预防和控制措施已成为工程应用的热点和难点。本发明提供一种数据降维与支持向量机的产品早期故障根原因识别方法,首先考虑到大数据背景下,产品寿命周期数据的高维数据信息导致分析困难,为了排除高维数据特征空间的数据冗余特征和噪声特性对分析结果的影响,利用主成分分析方法进行数据降维。其次,考虑到传统对故障认识的分散性,本发明针对早期故障特征结合降维后的数据,借鉴公理化域映射与瀑布式分解理论系统的建立起故障关联树。最后,利用支持向量机技术分类确定出不同层级间影响早期故障发生的根原因。本发明形成了大数据环境下对早期故障特性进行高维特性降维建树的过程,建立起产品早期故障特征与各级影响因素间的映射关系并确定回溯问题的根本所在。本发明从产品可靠性形成角度大数据出发,从根本上弥补传统意义上忽略高维数据特征空间而导致的对早期故障机理认识的不精确与误判。

[0007] (2) 技术方案:

[0008] 本发明是一种基于降维与支持向量机的产品早期故障根原因识别方法,提出的基本假设如下:

[0009] 假设1可靠性设计方案不变。

[0010] 假设2产品设计、制造到使用环节过程是可测量的。

[0011] 假设3过程检测值相互独立。

[0012] 假设4销售、运输环节对产品早期故障的影响不考虑。

[0013] 基于上述假设,本发明提出的一种基于降维与支持向量机的产品早期故障根原因识别方法,其特征在于:其步骤如下:

[0014] 步骤一 构建故障特征导向的关联树概念模型;

[0015] 步骤二 构建故障根原因节点大数据模型;

[0016] 步骤三 产品寿命周期质量与可靠性数据收集;

[0017] 步骤四 构建主成分分析技术模型;

[0018] 步骤五 基于主成分得分的关联树构建;

[0019] 步骤六 构建支持向量机技术分类模型;

[0020] 步骤七 基于支持向量机技术分类的节点优先级排序；

[0021] 步骤八 结果分析。

[0022] 其中，步骤一中所述的“构建故障特征导向的关联树概念模型”，其构建的做法如下：

[0023] 借鉴公理化域映射理论与瀑布式分解理论，从系统论角度自上而下将早期故障特征映射到功能域，实现功能分解；进而，每个功能需求转化到物理域完成物理分解，继而，被映射到过程域，完成工艺分解并最终形成故障特征导向的关联树概念模型。

[0024] 其中，步骤二中所述的“故障根原因”是关联树中的关键节点，是指区别于直接原因且具有潜在性的特点，是导致故障发生的最根本的原因，。

[0025] 其中，步骤二中所述的“构建故障根原因节点大数据模型”，构建方法如下：

[0026] 首先，针对目前对早期故障机理的笼统归因于设计、制造和使用无法明确定位的问题，立足于产品早期故障症状，从产品最终可靠性形成角度出发，从设计、制造和使用进行系统梳理导致产品早期故障的原因相，提取相关数据；最后，依据所提取的数据信息，建立起潜在故障根原因大数据模型。

[0027] 其中，步骤三中所述的“产品寿命周期质量与可靠性数据收集”是指利用传统历史数据和专家经验等手段基于步骤2的大数据模型收集设计、制造、使用三方面的为主的产品生命周期内的质量检测历史数据以及故障数据。

[0028] 其中，步骤四中所述的“构建主成分分析技术模型”，其构建方法如下：

[0029] 首先，在关联树概念模型中选择进行数据降维的目标节点。其次，针对建立的故障根原因大数据数据模型，将原始数据集的多个维度等进行整理，作为训练样本。

[0030] 给定一个d维空间的具有n个过程变量的N个数据样本， $X = [x_1, x_2, \dots, x_n]$ 是原始故障数据输入向量集， $\bar{x}$ 是样本均值，其协方差计算为：

$$[0031] \quad R = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(x_i - \bar{x})^T$$

[0032] 给定一个n维特征向量 x_i ($i=1, 2, \dots, m$)，特征向量V R的正交投影方向为，

$$[0033] \quad RV = \Lambda V$$

[0034] 其中 Λ 是特征值，特征值能通过解下面的方程获得。

$$[0035] \quad |R - \lambda I| = 0$$

[0036] 其中 λ ($i=1, 2, \dots, n$)是特征值，I是单位矩阵。

$$[0037] \quad \Lambda = \text{diag}[\lambda_1, \lambda_2, \dots, \lambda_k]$$

[0038] 继而，通过变换把原始的数据通过该变换，变换到一个新的空间中，在这个新的变换空间，能够找到原始数据主要的属性，通过对这些属性进行选取前d个分量降到d维空间，降低数据的维度。

[0039] 其中，步骤五中所述的“基于主成分得分的关联树构建”，其构建方法如下：

[0040] 把原始数据进行数据的标准化处理，各PCA提取主成分，得到各主成分下的新数据，即为主成分所占的分值，再利用主成分得分值确定各个节点建立故障关联树，其中主成分得分的具体形式如下：

$$[0041] \quad \begin{pmatrix} F_{11} & F_{12} & \cdots & F_{1k} \\ F_{21} & F_{22} & \cdots & F_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ F_{n1} & F_{n2} & \cdots & F_{nk} \end{pmatrix}$$

[0042] 其中,步骤六中所述的“构建支持向量机技术分类模型”,其构建方法如下:

[0043] 支持向量机技术是基于结构风险最小化理论之上在特征空间中建构最优分割超平面。假设输入向量集为 $T = (x_i, y_i)$,构造最优超平面的问题转化为下列二次规划问题。

$$[0044] \quad \min_{w, b} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \xi_i$$

$$[0045] \quad \text{St. } y_i (w^T x_i + b) \geq 1 - \xi_i, \xi_i \geq 0$$

[0046] 这里 w 和 b 为因子, ξ_i 是松弛变量, C 为惩罚因子。该最优化问题可以转化为应新的二元问题为

$$[0047] \quad \max_{\alpha} \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N y_i y_j \alpha_i \alpha_j x_i^T x_j$$

$$[0048] \quad \sum_{i=1}^N \alpha_i y_i = 0, \alpha_i \geq 0$$

[0049] 在特征空间,为了使最大化边界 $2/\|w\|$ 找到最优超平面,决策函数转换为,

$$[0050] \quad f(x) = \text{sgn} \left\{ \sum_{i=1}^N \alpha_i y_i K(x, x_i) + b \right\}$$

[0051] 这里 x 表示识别的样本, b 为阈值, α_i 为拉格朗日乘数, $\text{sgn}(u)$ 为符号函数, $K(x_i, x_j)$ 为核函数,选择高斯径向基函数为核函数,表达如下:

$$[0052] \quad K(x, x_i) = \exp(-\gamma \|x - x_i\|^2)$$

[0053] γ ($\gamma > 0$) 为高斯核函数宽度,最终问题可转换为如下方程:

$$[0054] \quad \max_{\alpha_i} \left\{ \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j K(x_i, x_j) \right\}$$

[0055] 本发明,利用支持向量机处理多分类数据,因此一对多的分类方法需要建立 $K(K-1)/2$ 个分类器。

[0056] 其中,步骤七中所述的“基于支持向量机技术分类的节点优先级排序”,是针对步骤5中估算的关联树节点主成分构建的故障关联树,利用步骤6构建的支持向量机模型进行故障关联树节点的分类。首先选择故障关联树中的目标节点;再根据支持向量机模型给定不同参数 C 和 γ 计算支持向量机分类器的准确率;最后,依据分类的准确率大小得到节点优先级排序。

[0057] 其中,步骤八中所述的“结果分析”,是指依据步骤7的结果分析对比不同参数下分类器的结果,得到故障关联树节点权重优先级,并最终确定产品早期故障的根原因。

[0058] 通过以上步骤,实现了产品寿命周期高维大数据环境下的故障根原因的溯源分析,在工程应用中为制造商实施积极的早期故障管控策略提供了明确的目标,可避免对早期故障原因的误判分析和控制的无效,有利于产品在设计、工艺设计阶段采取事前预防和

整改措施。

[0059] (3) 本发明所述的一种基于降维与支持向量机技术的产品早期故障根原因识别方法,其使用方法如下

[0060] 步骤1构建故障特征导向的关联树概念模型;利用公理化域映射与瀑布式分解理论构建早期故障特征导向的关联树概念模型,在此过程中不考虑运输和使用的情况下,构建从设计到制造的功能域、物理域、工艺域的故障关联树概念模型。

[0061] 步骤2构建故障根原因节点大数据模型;依据早期故障内涵,从产品可靠性最终形成出发,建立起以早期故障为目的,制造过程为核心的,上游追溯到设计、下游立足于使用的大数据模型。

[0062] 步骤3从产品最终可靠性形成角度出发,收集产品寿命周期质量与可靠性数据。

[0063] 步骤4基于构建的故障关联树概念模型,从针对一个具体故障症状如何从海量数据中找到主节点构建起故障关联树出发,考虑到降维的必要性,建立起主成分分析技术模型。

[0064] 步骤5根据主成分分析技术模型,估算节点的主成分得分,即为故障关联树各层级的节点,完成故障特征导向的关联树的构建。

[0065] 步骤6针对故障关联树的节点,构建支持向量机技术分类模型辅助分析。

[0066] 步骤7基于步骤6建立的支持向量机技术模型,利用一对多分类器的准确率得到节点相对效率值,并对得到的节点优先级排序。

[0067] 步骤8讨论分析节点权重,针对关联树节点的分类准确率得到产品早期故障的根原因。

[0068] (4) 优点和功效:

[0069] 本发明是一种基于降维与支持向量机的产品早期故障根原因识别方法,其优点是:

[0070] i. 本发明提出的基于降维的故障特征导向的故障关联树概念模型,突破了产品寿命周期环境下高维大数据的噪声和冗余导致的故障关联树的偏差构建,利用主成分分析进行降维处理有效的提取了数据信息,是数据降维的预处理过程是准确进行故障根原因识别不可或缺的部分。

[0071] ii. 本发明利用主成分分析技术和支持向量机技术去分类识别故障关联树节点,定位关键节点,实现了故障根原因的溯源分析,可避免对早期故障原因的误判分析和控制的无效,有利于产品在产品设计、工艺设计阶段采取预防和整改措施。

附图说明

[0072] 图1是本发明所述方法流程图。

[0073] 图2是早期故障特征导向的根原因关联树概念模型。

[0074] 图3是早期故障特征导向的根原因关联树。

[0075] 图中符号说明如下:

[0076] SVM2 (PV1.3.2) 是指分类器即针对关联树节点 (过程域中的PV1.3.2)

具体实施方式

[0077] 下面将结合附图和实例对本发明做进一步详细说明。

[0078] 本发明一种基于降维与支持向量机的产品早期故障根原因识别方法,见图1所示,其具体步骤如下:

[0079] 步骤一、构建故障特征导向的关联树概念模型;

[0080] 构建关联树的目标是从产品最终可靠性形成角度出发,利用公理化域映射和瀑布式分解理论系统的搭建起一个贯穿设计到制造的可能导致早期故障发生的各个环节。通过构建的关联树概念模型在产品功能域、产品物理域和产品过程域中找可能的节点,最终实现故障关联树概念模型如图2所示。

[0081] 步骤二、构建故障根原因节点大数据模型;

[0082] 产品寿命周期通常包括设计、制造、使用、维修和回收环节。在设计阶段,顾客需求转化成详细的设计方案,然后设计方案经生产制造阶段得以实现。在使用阶段,一般顾客使用产品时会有制造商提供相应的保修服务。如果在早期使用阶段出现故障,产品进入维修阶段,直至不能再使用,它将寿命终结进入回收和废弃阶段。因此,为了追溯产品寿命周期的相关潜在因素,主要任务就是从设计、制造和使用中找到相关大数据。

[0083] 步骤三、产品寿命周期质量与可靠性数据收集;

[0084] 为了在早期故障发生的情况下系统化的找到潜在原因,从包括设计、制造、测试和使用的整个产品寿命周期收集质量和可靠性相关的大数据是必不可少的也是很有价值得数据库资源。实例中的数据来源于产品寿命周期各个环节,来自于人、机、料、法、环、测的偏差及不确定因素会导致关键特性产生波动与偏差,并最终影响了产品质量与使用初期产品的可靠性。在这些不确定的因素中,选取了其中的12个因素(温度,环境,可检测度,故障概率,故障严酷度等)进行产品故障根原因分析,由于制造过程是影响产品早期故障的主要环节,因此实例部分主要以产品电路板制造环节的变压器(PV1.3)为例,如表1所示为电路板故障潜在因素表,使用matlab模拟产生产品生命周期中1800条事务数据记录。

[0085] 表1洗衣机箱体噪声故障的潜在故障因素

过程节点				
FR1(电路调节)	FR2(电源)	FR3(供电保护)	...	
DP1.1(热敏电阻)	DP1.2	DP1.3	...	
	(二极管)	(变压器)		
PV1.3.1(剪裁)	PV1.3.2(卷绕)	PV1.3.3(退火)	PV1.3.4(线圈装配)	...
PV1.3.4.1(引线焊接)	PV1.3.4.2(浇注)	PV1.3.4.3(缠绕)	PV1.3.4.4(装配工艺)	...
...	...	...	...	...

[0087] 步骤四、构建主成分分析技术模型；

[0088] 由于产品生命周期中数据分析的维度太大，产生的数据量也很大，直接对数据进行分析，不仅会带来较大的时间成本，也不易于产品故障根原因的分析，因此采用主成分分析技术对数据进行降维，在保留数据主要成分的同时减少对数据的分析量。

[0089] 案例中将12个维度的数据通过PCA降维后，得到一个新的4个维度，即将该阶段产品可能产生故障4个子环节，剪裁 (PV1.3.1)、缠绕 (PV1.3.2)、退火 (PV1.3.3) 和线圈装配 (PV1.3.4) 四大分类的数据，保留了原有数据中86.22%的属性。通过归一化原始数据，求得相应的协方差矩阵和特征值和特征向量如下表。

[0090] 通过PCA降维最终得到的新数据，其部分数据表如下：

[0091] 表2数据表

[0092]

Number	1	2	3	4
1	1.4774	-0.5773	0.7983	1.7282
2	1.6409	-1.3440	0.8323	0.9957
3	3.9624	0.8990	1.5626	0.1806
...	...	...	...	...
1800	1.4508	0.0973	-0.6946	-1.3212

[0093] 步骤五、基于主成分得分的关联树构建；

[0094] 利用上述主成分分析技术将设计域，物理域与工艺域中所有对象降维提取出来，建立如图3所示的故障根原因关联树模型。

[0095] 步骤六、构建支持向量机技术分类模型；

[0096] 通过对PCA降维得到的新数据Z进行分析，选取70%的数据作为训练数据，得到训练模型，基于得到的模型对剩下的30%的数据进行分类预测，并计算相应的准确率。案例基于高斯核函数，对产品PV1.3阶段产生故障的主要原因PV1.3.1、PV1.3.2、PV1.3.3、PV1.3.4分别构造了SVM1、SVM2、SVM3、SVM4四个分类器。

[0097] 步骤七、基于支持向量机技术分类的节点优先级排序；

[0098] 针对SVM2 (PV1.3.2) 分类器，通过在给定不同的参数C和 γ 的情况下，得到SVM2分类器的准确率为如下表3所示：

[0099]

准确度		γ			
		1	0.5	0.25	0.125
C	1	0.9194	0.9024	0.9194	0.9294
	2	0.8698	0.8598	0.8998	0.8498
	4	0.9166	0.9139	0.9111	0.9139
	8	0.8340	0.8450	0.8450	0.8395

[0100] 根据上表中不同的参数C和 γ 得到最优的SVM2 (PV1.3.2) 分类器的准确率为：92.94%，此时C=1, $\gamma=0.125$ 。

[0101] 根据同样的方法，可以得到分类器SVM1、SVM3、SVM4在不同参数下的最优分类器及相应的准确率，实验结果如下表所示：

[0102] 表4实验结果

[0103]

分类器	C	γ	准确率
SVM1	1	0.5	86.64%
SVM2	1	0.125	92.24%
SVM3	4	0.125	87.76%
SVM4	1	0.125	90.27%

[0104] 步骤八、结果分析

[0105] 根据上表4中SVM分类器的准确率可以得到，在产品PV1.3阶段中，准确率排序即为优先级排序：PV1.3.2>PV1.3.4>PV1.3.3>PV1.3.1，因此，PV1.3.2是导致电脑板故障的主要原因。

[0106] 同样的，针对所有的功能和制造环节，利用上述方法最终能得到关键的设计参数和关键过程变量，它们将是洗衣机箱体早期故障预防的重要环节，是对早期故障率管控的重点。洗衣机箱体电脑板可靠性的形成过程，可发现其复杂的结构特性及数目众多的组件及零件等，增加了制造过程控制的难度。过程质量波动如芯片含非致命性缺陷，组件不合格，装配失调等不可避免的存在于电脑板的制造过程，带来制造缺陷，严重影响了被制造产品的固有可靠性，导致较高的早期故障率。因此，展开制造过程波动的偏差效应对早期故障率的影响分析，有利于早在制造阶段就能对早期故障率进行准确预计并展开预防性的控制。目前针对工业4.0智能制造模式对于质量与可靠性保障的新需求，基于大数据的制造产品早期故障防控成为亟待解决的问题，传统对早期故障的认识缺乏对制造过程数据的处理，智能制造时代使得制造过程数据的获取成为可能，本专利出于可靠性保障目的，以高维大数据特征为防控对象，利用本专利所提的方法对高维数据进行降维处理再定位分析这些典型故障症状的根原因，有助于后续采取相应措施去管控产品实现过程中的关键薄弱点，进而避免发生类似的质量问题，在工程应用中为制造商提供了明确的目标去实施积极的早

期故障管控策略。

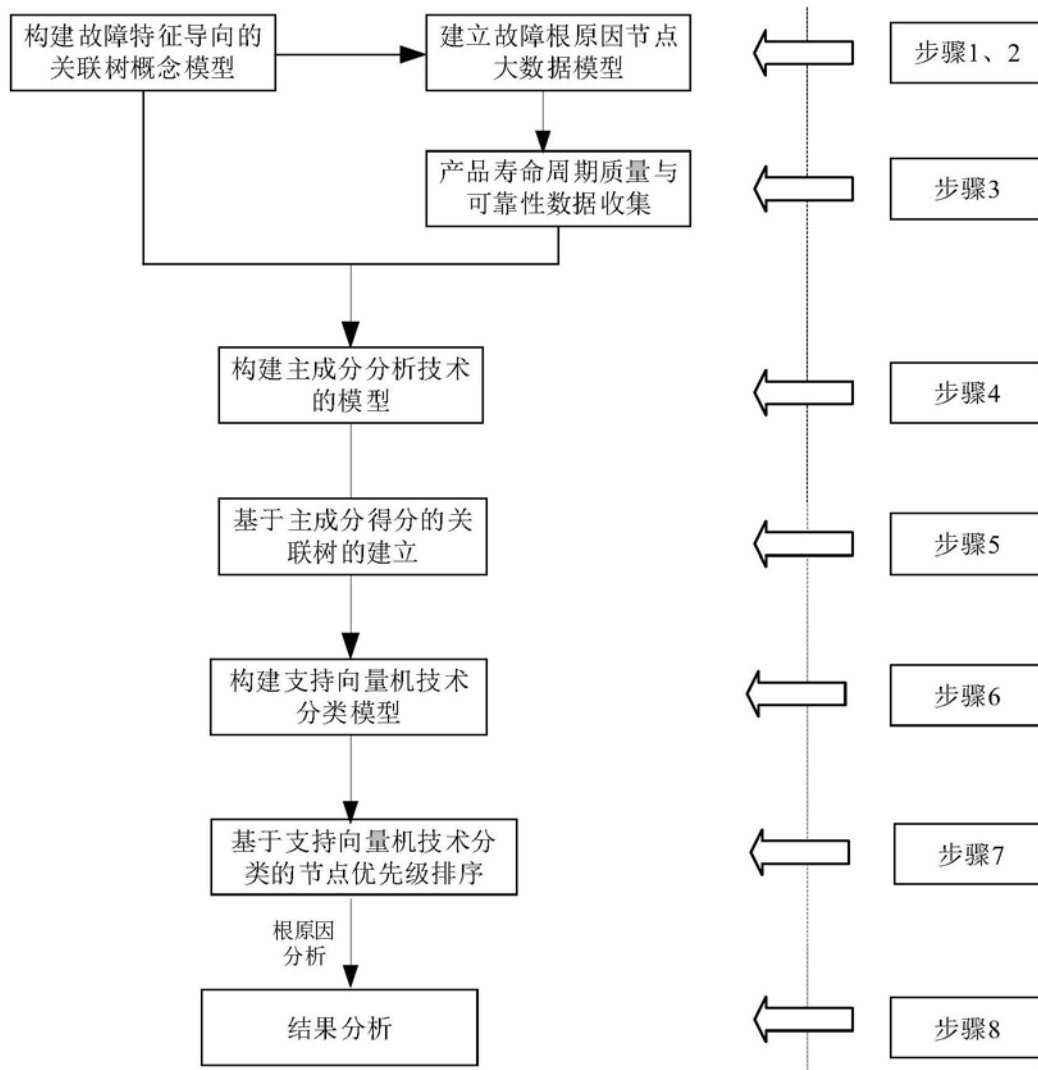


图1

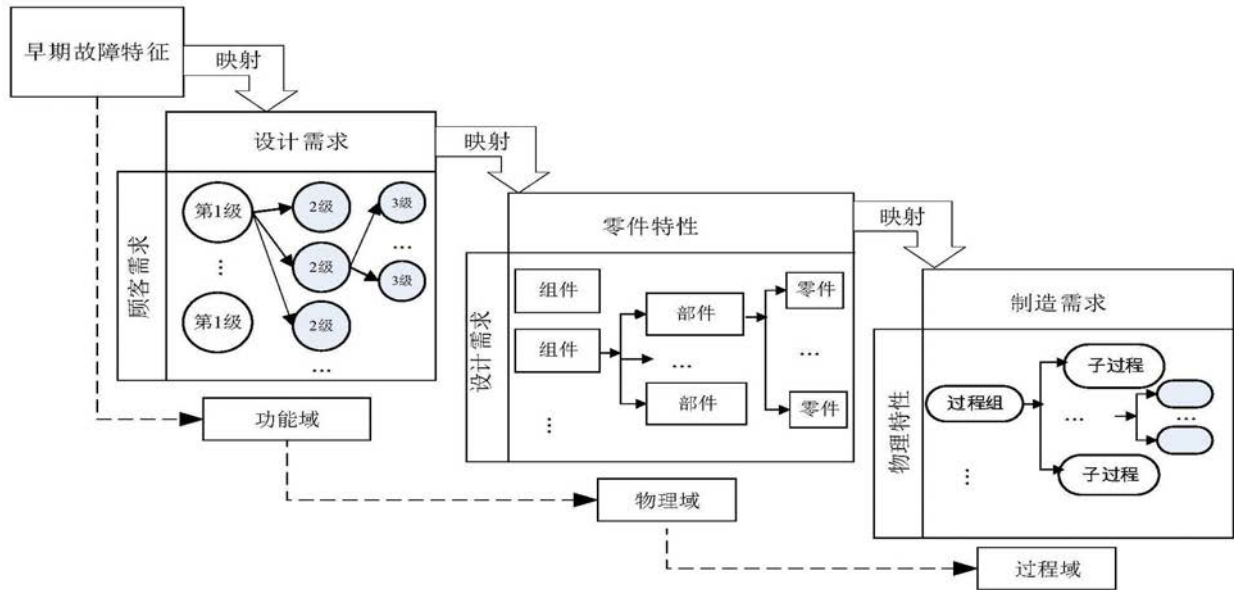


图2

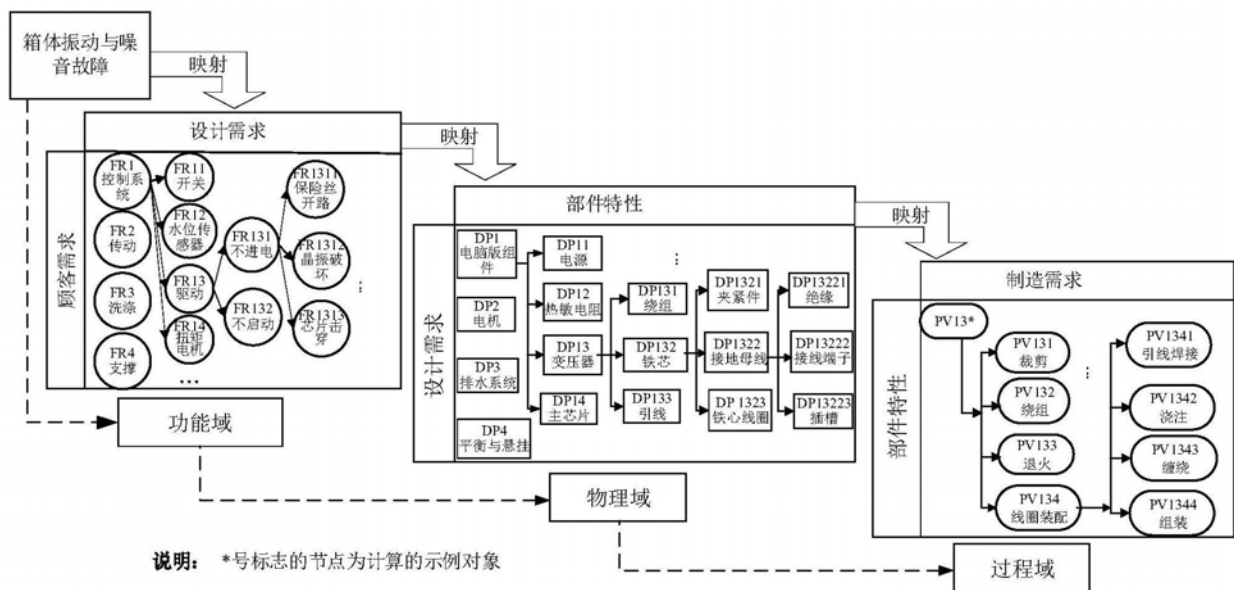


图3