DOI:10.20033/j.1003-7241.(2025)02-0061-05.

基于深度学习算法的物联网资源信息自动采集方法

严 梅

(浙江机电职业技术学院,浙江 杭州 310053)

摘 要:常规物联网资源信息自动采集方法,多采用本体构建技术,无法对其中的冗余信息进行有效滤除,导致资源信息采集效率较低。因此,提出基于深度学习算法的物联网资源信息自动采集方法。根据物联网运行特点设定物联网资源信息采集周期,控制采集过程,利用卡尔曼滤波法清洗原始数据,滤除冗余信息,提取关键信息进行内部融合,基于信息融合结果对其特征向量进行提取,分析信息概率分布情况,利用深度学习算法扩展信息采集量的最大化范围,进而实现物联网资源信息自动采集。采用仿真实验的方式验证所提方法的应用性能,结果证明,所提方法具有更高的资源信息采集效率。

关键词:深度学习算法;物联网;资源信息;自动采集

中图分类号:TP391 文献标识码:A 文章编号:1003-7241(2025)02-0061-05

Automatic Collection Method of Internet of Things Resource Information Based on Deep Learning

YAN Mei

(ZheJiang Institute of Mechanical & Electrical Engineering, Hangzhou 310053 China)

Abstract: Conventional IoT resource information automatic collection methods mostly use ontology construction technology, which can not effectively filter redundant information, resulting in low resource information collection efficiency. Therefore, an automatic collection method of IoT resource information based on deep learning algorithm is proposed. According to the operation characteristics of the Internet of Things, it sets the collection cycle of the resource information of the Internet of Things, controls the collection process, uses Kalman filtering to clean the original data, filters out redundant information, extracts key information for internal fusion, extracts its feature vector based on the information fusion results, analyze the probability distribution of information, uses the depth learning algorithm to expand the maximum range of information collection, and then realizes the automatic collection of the resource information of the Internet of Things. The application performance of the proposed method is verified by simulation experiments, and the results show that the proposed method has higher resource information collection efficiency.

Keywords: deep learning; Internet of Things; resource information; automatic acquisition

0 引言

根据物联网运行特点对物联网资源数据进行自动采集是监测物联网运行状态的重要任务之一。随着该项工作的深入开展,相应的信息自动采集方法也越来越多。文献[1]基于时间序列的信息采集方法通过构建物联网环境数据和运行数据相互分离的时间窗信息序列采集模型,并结合数据挖掘技术提高信息采集精度;文献[2]基于5G边缘计算的信息采集方法主要利用5G技术计算物联网信息采集的时间判据,并采用辅助分析系统实现信息自动采集。但以上方法无法适应物联网资源信息公式的变动,存在信息采集效率低的弊端。本文针对上述问题进行了相关研究,利用深度学习算法对物联网资源信息采集方法展开优化,以此实现资源信息的自动采集,为物联网环境管理提供便利。

1 物联网资源信息自动采集方法设计

1.1 设定物联网资源信息采集周期

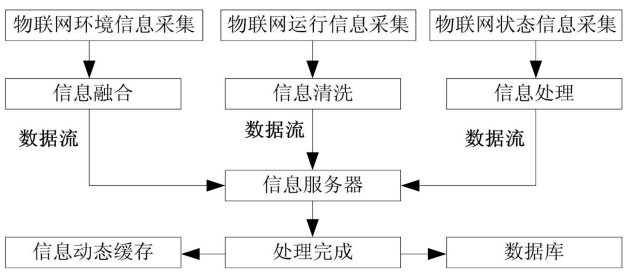


图1 周期性资源信息采集模式示意图

对网络资源信息自动采集的主要流程是利用技术手段获取物联网原始资源信息,然后对初始信息进行预处

*基金项目:浙江机电职业技术学院教育教学改革项目(S-0152-21-471);浙江机电职业技术学院项目(S-1001-21-300-04-110)

收稿日期:2023-08-07

理后完成信息采集工作。针对资源信息流向过程，为物联网设定资源信息采集周期，控制采集过程，具体采集过程如图1所示。

分析图1可知，物联网资源信息采集周期是一种采集模式，通过获取物联网环境信息、运行信息以及状态信息，经过信息融合、清洗和处理后传输到信息服务器中，最后将信息动态缓存到数据库中[3]。本文将此周期性信息采集模式作为资源信息采集方法的基础，设计出一种更加便捷和高效的信息采集方法，控制物联网的信息采集周期。此信息采集周期将依照物联网的运行状态进行调整，为此构建物联网运行状态预测模型，确定信息采集周期。

假定某一特定资源信息 A_{ij} 流向下一网络节点 W_{ijk} 的过程时间为 $\max\{B_{ij1},B_{ij2},\cdots,B_{ijn}\}$，起始时间为 T_{ijn}，且 W_{ijk} 与其相邻节点之间的通信长度为 T_k，则将信息 A_{ij} 流向所有网络节点的时长相加，即可得到资源信息 A_{ij} 的全部流动时间。由此可以确定：

若 W_{ijk}!=null，则信息流动的起始时间为：

$$W_{ijn}=W_{ijn}^{'} \tag{1}$$

式中，W_{ijn} 表示信息流动的起始时间；W'_{ijn} 表示信息对应的完备基向量。

若 W_{ijk}!=null，此时 A_{ij} 的流动还未开始。将已经采集到的有关于信息 A_{ij} 的最迟流动起始时间设定为 W_{ijn}，然后找到 W_{ijk} 与 W_{ijn} 之间的流动时长，并将其全部相加，得到此信息的流动预计初始时间，则有：

$$W_{ijk}=\max\left\{B_{ij1},B_{ij2},...,B_{ijn}\right\} \tag{2}$$

$$W_{ijn}^{'}=W_{ijn}+\sum_{i=j}^{n-1}T_{ijn}+W_{ijk} \tag{3}$$

式(2)中，B_{ijn} 表示信息 A_{ij} 在第 n 个传输信道中的功率增益；式(3)中，T_{ijn} 表示流动起始时间。

根据式(2)和式(3)可以得到资源信息向相邻网络节点流动的起始时间和结束时间，因此也就得知了信息的流动过程，接下来计算流动过程中的完成度，即：

$$E_{ijn}=\frac{D-W_{ijn}}{T_{ijn}\bullet W_{ijn}^{'}}*100\% \tag{4}$$

式中，D 表示当前时刻信息 A_{ij} 与下一通信节点之间的距离。

为保证信息采集周期的稳定性，基于式(4)控制物联网时钟质量[4]。假设物联网时钟信号每隔一定时间间隔振动一次，即：

$$v(t)=v*\sin(2\pi st) \tag{5}$$

式中，v 表示时钟最大输出振幅；s 表示实际频率。

根据上式，可得到时钟振动的斜率为：

$$\frac{d_v}{dt}|\max=2\pi s*Y*V+\sum E_{ijn} \tag{6}$$

式中，Y 表示时钟最大振幅与实际频率的比值。根据此公式，获取信息采集周期，以此确定信息采集过程：

$$D_{Yi}=\frac{Y}{\pi s*2^{n+1}}*\frac{d_v}{dt} \tag{7}$$

式中，D_{Yi} 表示信息采集周期。

使用此公式实现信息采集周期的设定，为后续资源信息融合打下基础，进而完成物联网资源信息的采集工作。

1.2 资源信息内部融合

根据物联网资源信息采集周期，获取物联网原始资源信息，但是原始信息中含有不可用信息，会对后续物联网关键信息的采集造成干扰，因此，选用卡尔曼滤波法清洗原始信息，滤除信息中的冗余信息，具体清洗过程如下：

信息处理结果：

$$H_i=Gr_i+o_i \tag{8}$$

式中，H_i 表示信息经过处理后的集合；G 表示 i 时刻时信息的转移矩阵；r_i 表示 i 时刻信息中不含噪声的信息量；o_i 表示信息的高斯分布。

将式(8)输入到采集周期中信息流动过程完成度公式中，对边缘信息进行清洗和滤除，提取关键信息，并对信息进行进一步融合[5]。假设采集到的物联网原始资源信息服从正态分布，其特征为(α,β)，其中α表示信息方差，β表示信息标准差。对于物联网，要保证信息采集误差不超过 ±0.5，设定物联网读写器个数为 N，采集到的原始信息总量为 U_i，则 J 个信息融合结果 X 可表示为：

$$X=D_{Yi}*\frac{\sum_{i=1}^{J}U_i}{J*H_i} \tag{9}$$

式中，D_{Yi} 表示信息采集周期；H_i 表示信息经过处理后的集合。

使用上述公式，实现信息采集结果的内部融合，而后根据融合结果确定物联网的关键信息，并对关键资源信息的特征向量进行提取，提高关键信息的采集效率[6]。

1.3 提取物联网资源信息特征向量

确定物联网关键资源信息后，接下来对信息的特征向量进行提取。假定 s'_{sfg} 表示信息特征向量的平均值，$\tilde{A}b$ 表示信息特征向量的局部标准差，$\hat{A}b$ 表示信息特征向量的局部期望值，ξ 表示能见函数，则可以通过下式计算出信息频率的边缘分布情况：

$$E'=X\bullet\frac{\xi\times s_{sfg}^{'}}{p^{'}\pm\tilde{A}b}\mp\frac{\left\{\hat{A}b\right\}}{e_1\oplus e_2} \tag{10}$$

式中，p' 表示固定参数；X 表示信息融合结果；e_1 表示首次采集信息所需要的时间；e_2 表示最后一次采集信息所需要的时间。

根据物联网资源信息的边缘频率分布情况，计算原始信息中不可用信息在学习中的分布密度为：

$$u=\sum E^{'}*\frac{k_{li}u_{fj}}{\chi_{fj}}\mp\frac{\left\{e_{sg}\times\xi_{wer}\right\}}{e_{sgh}\times k_{li}} \tag{11}$$

式中，k_{li}表示网络节点之间的建立权重；u_{fj}表示物联网当前环境状态；χ_{fj}表示学习动作空间；e_{sg}表示学习样本数量的估计值；ξ_{wer}表示网络节点间的最短通信路径；e_{sgh}表示信息状态分布概率。

利用滤波器对物联网资源信息的参数进行调整，调整公式为：

$$E_{ue}=\frac{\mu_{op}\times t_{sj}}{u}\mp\frac{\left\{g_{we}\otimes s_{sgj}\right\}}{f_{ert}} \tag{12}$$

式中，μ_{op}表示信息变化率系数；t_{sj}表示信息阈值；g_{we}表示不包含冗余信息时的物联网状态；s_{sgj}表示信息距离矩阵；f_{ert}表示特征向量提取的门限值。

基于式(11)的调整结果，提取出物联网资源信息的特征向量，表示为：

$$s=\frac{\sigma_{qwe}\mp k_{wer}}{E_{ue}\oplus u} \tag{13}$$

式中，σ_{qwe}表示信息特征的局部幅度平滑性；k_{wer}表示信息特征的矢量空间；E_{ue}表示信息参数调整结果；u表示冗余信息密度。

以上，根据时域理论，得到了物联网资源信息的边缘频率分布情况，调整信息的参数，提取出信息的特征向量，为物联网资源信息自动采集提供了有利条件[7]。

1.4 利用深度学习算法采集物联网资源信息

深度学习算法是指从一个信息特征向量中产生一个输出结果所涉及的计算可以通过一个流程图来表示。基于资源信息特征向量提取结果，在此利用深度学习中的神经网络算法实现物联网资源信息的自动采集工作。具体实现流程如图2所示。

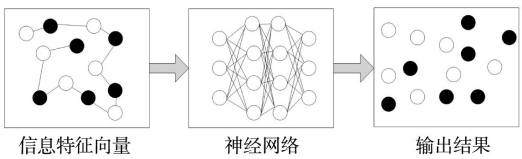


图2 利用神经网络自动采集的物联网资源信息流程

图2中一般包含两类信息，即信息节点的属性信息与结构信息。前者主要是用于表示信息节点所代表的对象的特性，而后者则侧重于描述对象与对象之间的相互作用关系[8]。

根据物联网特点，计算出实际发送信息的总量，公式为：

$$q=s\times\frac{q_{total}\times e_{dh}}{d_t d_s}\pm t_{sh}\times t_{ssi} \tag{14}$$

式中，s表示信息特征向量；q_{total}表示信息采集总量；e_{dh}表示噪声干扰检测值；d_t表示信息时间矢量数据；d_s表示信息传输速率；t_{sh}表示信息中的无用数据；t_{ssi}信息表示空间矢量数据。

在上式基础上，扩大信息采集最大化范围，扩展条件为：

$$r_y=\frac{\varsigma_d\mp h_s}{p_w\mp q}\otimes f_{ip} \tag{15}$$

式中，ς_d表示信息发送端的天线增益；h_s表示信息接收端的天线增益；p_w表示信息通信链路损耗；q表示物联网实际发送信息量；f_{ip}表示天线功率。

设定p_r表示物联网节点间能量损耗因子，k_{dh}表示物联网载波波长，p_{ery}表示信息采集速率，得到物联网通信节点接收信息总量与跳数之间的关系，表示为：

$$E^{''}=\frac{p_r\pm k_{dh}}{p_{ery}}\oplus\frac{E^{'}_{per}\mp d^{'}_{sh}}{y_{ER}\mp r_y} \tag{16}$$

式中，E'_{per}表示节点总发射能量；$d^{'}_{sh}$表示相邻两个节点距离的平方；y_{ER}表示信息链路形成的最长路径跳数；r_y表示扩展条件。

根据物联网节点接收信息总量[9]与跳数之间的关系，给出信息采集的目标函数，即：

$$r_{gi}=\frac{p^{'}\pm k_{xj}}{u_{dgh}+l_{dk}}\times E^{''} \tag{17}$$

式中，$p^{'}$表示网络节点稀疏度；k_{xj}表示信息采样频率；u_{dgh}表示信息序列平均值；l_{dk}表示信息采集时间间隔。

基于以上过程，得到物联网资源信息的采集结果为：

$$F=\frac{y_{ER}\times r_{gi}}{f_{ip}}\mp u_{dgh} \tag{18}$$

式中，y_{ER}表示信息的物理量；r_{gi}表示信息采集的目标函数；f_{ip}表示最短跳数；u_{dgh}表示采集时间间隔。

通过以上分析与计算过程，确定网络节点发送与接收的信息总量，利用采集范围最大化扩展条件及给定信息采集目标函数，最终实现物联网资源信息的自动采集[9]。

2 实验

采用对比实验的方法，对所提方法进行仿真验证和分析，测试设计方法的可行性。

2.1 实验准备

在物联网100 m×100 m的监测区域部署N(1 024)个网络从节点和一个传感器节点，这些从节点服从随机均匀分布，并且它们的位置已知[10]。传感器节点部署在监测区域的(50,50)位置，它主要负责压缩信息的采集及信息的恢复。网络中任意两个节点之间的通信半径为R，它的初始值设置为$0.1\times d_{max}$m，此时网络连通。

将目标监测区域划分为基础网络、局部网络和信息采集服务器网络三部分。由于物联网的网络节点较多，且较为复杂，因此，模拟实际物联网结构，对物联网资源信息进行周期性采集[11]。设置的仿真实验参数如表1所示。

表1 仿真参数设置

项目	参数
节点数	100、200、300、400
路由协议	Stationary,Peripheral:AODV、DCSR、Random
节点初始能量	1.8 J
信息采集速度	401 kB/ 10 s
存储扩展方式	SD扩展
Sink 移动速度	Stationary: 0 m/s,其他协议:5 m/s,25 m/s
时间区间长度	2 s
仿真范围	100 mx100 m

在以上仿真实验传输设定下，确定物联网资源信息自动采集过程包括信息分类、信息解读、信息传输以及信息接收4个环节。采用Flask作为实验开展的核心框架，网络通信节点的交互服务器为HTTP客户端，同时显示信息采集状态[12]。设计的网络参数为：信息采集总时长3 600 s，网络资源节点个数为10，节点发送和接收信息量的功率增益为50 dB，采样频率为100 kHz，采样间隔为2 s。

2.2 实验说明

基于上述实验环境，采集网络节点初始信息，得到的采集信息时间序列如图3所示。

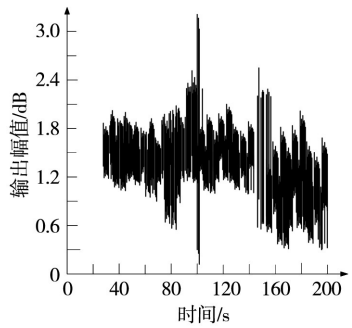


图3 信息采集时间序列

如图3所示，在采集时间为100 s时，信息的时间序列的输出幅值最大，表明在此之后的信息已进入持久化状态，并存入数据库。

仿真实验环境主要为物联网资源信息采集方法的应用提供基础支持，并与实验对照组，参照实验给定的结果评价指标对不同信息采集方法的实际性能进行检定。

2.3 信息汇聚量实验分析

使用设计的方法和文献[1]基于时间序列的信息采集方法(方法1)、文献[2]基于5G边缘计算的信息采集方法(方法2)在同一实验环境设置下，对物联网资源信息进行采集，比较不同方法的采集效率。得到的对比结果如图4所示。

根据图4可知，在10次实验中，方法1的平均信息汇聚量大约为330条，其中最大汇聚量为420条，可见其信息采集效率最低，主要成因是方法将信息传输到指定传感器中，然后融合所有传感器的信息，因此，工作时间相对较长，采集效率较低；方法2的平均信息汇聚量大约为590条，但该方法忽略了节点访问路线的差异，因此，采集效率还有待提高；本文方法的平均信息汇聚量在1 080条，采集效率最高。由此可以说明，本文设计的方法应用于物联网资源信息采集中，具有较高的信息采集效率。

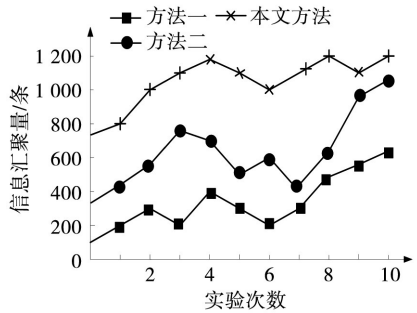


图4 不同信息采集方法信息汇聚量对比结果

2.4 资源信息采集速度实验对比

结合以上实验结果，为进一步验证设计的方法的信息采集高效性，对三种信息采集方法的信息采集速度进行实验，对比结果如图5所示。

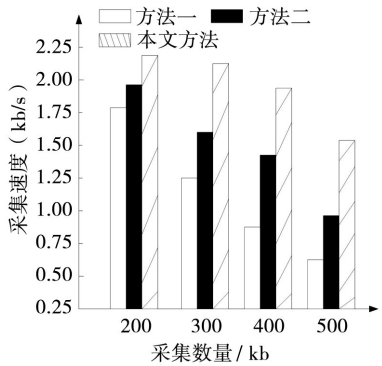


图5 三种方法资源信息采集速度对比结果

如图5所示，三种方法的信息采集速度都随着采集数量的增加而有不同程度的下降，但本文设计的方法的整体采集速度明显高于其他两种方法，因此，所提方法的信息采集效率更高。

3 结束语

本文研究了物联网资源信息采集场景，并提出了基于深度学习的物联网资源信息自动采集方法，通过设定采集周期、提取信息特征向量以及信息融合，从而保证完成资源信息采集任务的效率。仿真结果表明，利用所提方法进行物联网资源信息采集，其采集效率更高。该方法还可以应用在同类型的环境中，具有较高的推广价值。

参考文献：

[1] 王婧骅，张娟，赵婉茹，等.基于时间序列的分布式光伏电站发电信息采集方法[J].电网与清洁能源，2022，38(6)：137-142.

[2] 刘圣林，龙永树.基于5G边缘计算的工业信息采集方案[J].

中国仪器仪表,2020(9):70-73.

[3] 王小珍.网络视听节目监管系统信息采集关键技术研究[J].传播与版权,2022(7):62-64.

[4] 屈茹,杨麒龙.基于无人机的无线传感器网络信息采集研究[J].产业创新研究,2022(18):120-122.

[5] 韩忠辉.基于SDN架构的海上试验场信息采集系统网络设计研究[J].电子设计工程,2022,30(11):113-116.

[6] 马浩铭.网站资源信息自动采集技术[J].湖北农机化,2020(12):27-28.

[7] 李岚清,王恒,晏晓峰.网络资源自动采集技术研究[J].电子元器件与信息技术,2020,4(5):56-58.

[8] 王鹏.基于UAV-WSN的空中信息采集[J].现代工业经济和信息化,2022,12(3):157-159.

[9] 吴博皓.基于Web的智能信息采集和发现系统研究[D].武汉:华中科技大学,2020.

[10] 严智,邹鑫灏,潘伟.基于无线网络的物联网数据传输质量监控系统设计[J].电子设计工程,2022,30(9):175-179.

[11] 胡珊.网络安全信息采集关键技术[J].电脑编程技巧与维护,2022(2):173-175.

[12] 李元涛.汽车厂焊装车间网络规划与信息采集[J].中国信息化,2022(1):71-72,70.

作者简介:严梅(1983-),女,硕士,讲师,研究方向:物联网应用开发。

(上接第3页)

图4中箭头所指为在0 kN载荷下转速1 880 r/min下外圈单点损伤为0.2 mm的理论外圈故障特征频率94.368 6 Hz及其倍频,可以发现,故障特征频率相对其他干扰信息更加明显,故障特征频率被有效提取,倍频成分较周围频率成分更突出。

图7中,箭头所指为在4 kN载荷下转速1 680 r/min内圈单点损伤为0.2 mm的理论内圈故障特征频率139.670 6 Hz及其倍频,可以发现,故障特征频率相对其他干扰信息更加明显,故障特征频率被有效提取,倍频成分较周围频率成分更突出。

包络解调方法和CSES方法同样被用于对比分析。由图5、图8所示的故障信号包络谱可得,受背景噪声干扰,滚动轴承滚动体故障特征频率成分存在但不明显,二阶倍频成分相对突出,故障特征提取效果不好。

3 结束语

综上所述,在提取该滚动轴承内外圈故障时,包络解调分析方法难以有效提取故障特征。传统方法和本文提出方法均能诊断出故障,但本文提出方法的故障特征频率在其强化包络谱中更加明显,诊断效果更好。

参考文献:

[1] Huang N E,Shen Z,Long S R,et al.The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis[J].Proceedings Mathematical Physical&Engineering Sciences,1998,454(1971):903-995.

[2] Jian Z A,Ry B,Rxg A,et al.Performance enhancement of ensemble empirical mode decomposition[J].Mechanical Systems and Signal Processing,2010,24(7):2104-2123.

[3] 艾延廷,方妍,田晶.峭度准则EMD与空域相关结合的滚动轴承故障特征提取[J].机械设计与制造,2019(12):213-216.

[4] LI Yongbo,SI Shubin,LIU Zhiliang,et al.Review of local mean decomposition and its application in fault diagnosis of rotating machinery[J].Journal of Systems Engineering and Electronics,2019,30(4):799-814.

[5] 吕明珠,苏晓明,刘世勋,等.基于VMD-SVM的滚动轴承退化状态识别[J].机械设计与制造,2020(1):96-100.

[6] Tang G J,Wang X L.Parameter optimized variational mode decomposition method with application to incipient fault diagnosis of rolling bearing[J].Journal of Xi'an Jiaotong University,2015,49(5):73-81.

[7] 刘前进,高丙朋,宋振军,王维庆.基于参数优化变分模态分解的变速工况下轴承故障诊断[J].轴承,2022(8):71-78.

[8] 俞惠惠,郑近德,潘海洋,等.基于自适应变分模态提取的低速重载滚动轴承故障诊断方法[J].振动与冲击,2022,41(11):65-71,113.

[9] Zhang J,Huang H,Tao L,et al.Multiscale Fractures Characterization Based on Ant Colony Optimization and Two-Dimensional Variational Mode Decomposition[J].IEEE Journal of Selected Topics in Applied Earth Observations& Remote Sensing,2018:1-9.

[10] 谭亚红,史耀.完备变分模态分解和多传感器卷积神经网络的轴承故障诊断方法[J].机床与液压,2022,50(14):182-188.

[11] 方学宠,娄益凡,吴安定,等.基于SVMD与SLLE的机械设备齿轮箱故障诊断方法[J].机械与电子,2022,40(1):36-41,47.

[12] 左娟娟,朱红杰,杨继党,等.基于半监督机器学习的复杂电网连锁故障诊断方法[J].自动化技术与应用,2024,43(12):47-50,92.

[13] 聊方伦,周平.基于PCA的火电厂电气设备故障诊断方法[J].自动化技术与应用,2022,41(3):39-43.

[14] 康伟,朱永生,闫柯,等.基于CSES和MED的滚动轴承微弱故障特征提取[J].振动.测试与诊断,2021,41(4):660-666,827.

作者简介:李金城(1996-),男,硕士研究生,研究方向:故障诊断。

通信作者:张凯(1981-),男,博士,副教授,研究方向:旋转设备可靠性,振动及在线健康监测。