

基于数据挖掘的铁路信号设备故障自动诊断研究

张 硕

(北京中铁建电气化设计研究院有限公司, 北京 100043)

摘要: 铁路信号设备故障的类型多, 故障难以识别, 导致当前铁路信号设备故障诊断方法存在一定的弊端, 如错误诊断率高、诊断实时性差。为了获得更加理想的铁路信号设备故障结果, 设计了基于数据挖掘的铁路信号设备故障自动诊断方法。首先分析当前铁路信号设备故障诊断的研究进展, 找到引起铁路信号设备故障效果差的原因; 然后采集铁路信号设备状态信号, 提取铁路信号设备工作状态特征, 并引入数据挖掘技术对铁路信号设备故障变化特点进行建模, 构建铁路信号设备工作状态的分类器, 从而实现铁路信号设备故障诊断; 最后选择几种经典铁路信号设备故障诊断方法并进行对比测试。结果表明, 相对于经典方法, 该方法克服了当前铁路信号设备故障方法存在的局限性, 铁路信号设备故障诊断的准确性也得到了明显改善, 铁路信号设备故障时间明显缩短。

关键词: 铁路信号; 设备故障; 智能诊断; 数据挖掘; 特征提取

Research on automatic fault diagnosis of railway signal equipment based on Data Mining

ZHANG Shuo

((Beijing) China Railway Construction Electrification Design and Research Institute, Beijing 100043, China)

Abstract: There are many types of railway signal equipment faults, which are difficult to identify, which leads to some disadvantages of current railway signal equipment fault diagnosis methods, such as high fault misdiagnosis rate and poor real-time diagnosis. In order to obtain more ideal results of railway signal equipment faults, an automatic fault diagnosis method of railway signal equipment based on data mining is designed. Firstly, the research progress of fault diagnosis of railway signal equipment is analyzed to find out the cause of poor fault effect of railway signal equipment. Then, the state signal of railway signal equipment is collected to extract the working state characteristics of railway signal equipment, and the data mining technology is introduced to model the fault change characteristics of railway signal equipment, and a classifier of working state of railway signal equipment is constructed. In order to realize the fault diagnosis of railway signal equipment, several classical methods of fault diagnosis of railway signal equipment are selected and tested. The results show that compared with the classical methods, this method overcomes the limitations of current railway signal equipment fault methods, improves the accuracy of railway signal equipment fault diagnosis, and shortens the fault time of railway signal equipment.

Key words: railway signal; equipment fault; intelligent diagnosis; data mining; feature extraction

0 引言

近年来, 随着我国铁路高速的发展, 铁路

运营的范围越来越广, 铁路信号系统是确保铁路运输安全的关键。在实际应用中, 铁路设备

不可避免会发生故障，尤其是铁路信号设备，如何尽可能减少铁路信号设备故障对铁路运输安全的影响，已成为当前有亟待研究的重大课题^[1]。铁路信号设备故障诊断可以及时发现存在的问题，有助于建立故障决策机制，因此受到了人们的高度重视^[2-3]。

铁路信号设备结构很复杂，同时受到环境因素的影响，使铁路信号设备故障变化多样，且具有一定的随机性，最初通过维修人员对铁路信号设备故障进行分析和判断，这对维修人员的经验要求相当高，经常因经验不足出现铁路信号设备故障误判断现象，延误铁路信号设备维修时间，严重时可能导致铁路安全事故^[4]。随后出现基于专家系统的铁路信号设备故障诊断方法，通过建立铁路信号设备故障诊断的知识库，根据知识库的规则对铁路信号设备故障进行识别，但是这种方法的诊断结果与知识库的丰富度密切相关，使得铁路信号设备故障结果极不可靠^[5]。近年来，随着人工智能技术的发展，出现故障树的铁路信号设备故障方法和人工神经网络的铁路信号设备故障方法，它们的诊断结果要优于专家系统^[6-8]，但是同样存在一定的局限性，如诊断错误率高、诊断实时性差等^[9-10]。

为了获得更加理想的铁路信号设备故障结果，设计了基于数据挖掘的铁路信号设备故障自动诊断方法。结果表明，数据挖掘一定程度上解决了传统铁路信号设备故障方法存在的问题，铁路信号设备故障诊断结果整体性得到明显改善。

1 基于数据挖掘的铁路信号设备故障自动诊断方法

1.1 铁路信号设备故障信号的噪声去除

铁路信号设备状态数据由于受外界环境影响，通常存在一定的无效信息，它们以噪声的形式存在，本文引入小波分析法去掉铁路信号设备状态数据中的噪声。设小波母函数 $\psi(t)$ 的

非线性变换结果为 $\hat{\psi}(\omega)$ ，如果 $\hat{\psi}(\omega)$ 满足式(1)条件，那么 $\hat{\psi}(\omega)$ 为基本小波，即

$$C_{\psi} = \int_R \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < \infty \quad (1)$$

将 $\psi(t)$ 进行伸缩和平移后，可以产生小波序列 $\psi_{a,b}(t)$ ，具体为

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right) \quad a, b \in R, a \neq 0 \quad (2)$$

铁路信号设备状态数据可以描述为

$$s(t) = x(t) + n(t) \quad (3)$$

小波分析方法去噪的目的是将 $s(t)$ 中 $n(t)$ 去掉，得到真实铁路信号设备状态数据 $x(t)$ ，小波变换为

$$W_{2^j} s(k) = \frac{1}{2^j} \sum_{n=0}^{N-1} s(n) \psi^*(2^{-j}n - k) \quad (4)$$

式中， $W_{2^j} s(k)$ 为小波变换的系数。

由于式(4)计算过程十分复杂，因此采用递归实现，铁路信号设备状态数据小波分解具体可以表示如下

$$\begin{aligned} xf(j+1, k) &= xf(j, k)h(j, k) \\ Wf(j+1, k) &= xf(j, k)g(j, k) \end{aligned} \quad (5)$$

通过阈值法将噪声的小波变换系数设为0，然后通过铁路信号设备状态数据小波重构得到无噪的数据，具体为

$$\begin{aligned} xf(j-1, k) &= \\ xf(j, k)h_0(j, k) &+ Wf(j, k)g_0(j, k) \end{aligned} \quad (6)$$

1.2 铁路信号设备故障特征提取

采用经验模态分解法提取铁路信号设备故障诊断特征，具体步骤如下：

1) 采用经验模态分解法对铁路信号设备状态数据进行处理，产生 n 个相应模态函数 $c_i(t)$ ， $i=1, 2, \dots, n$ 。

2) 从铁路信号设备数据的模态函数分量得能量矩，具体如下

$$E_i = \int_{-\infty}^{+\infty} t |c_i(t)|^2 dt \quad (7)$$

通过式(7)可以得到铁路信号设备状态数据的能量矩 $E_1, E_2, \dots$ 。

3) 将多个能量矩组合在一起构造铁路信

号设备故障诊断特征向量 T ，具体如下

$$T = [E_1 \ E_2 \ \cdots \ E_n] \quad (8)$$

由于各个铁路信号设备故障诊断特征向量之间差异比较大，因此对其进行归一化操作，具体为

$$T' = [E_1/E \ E_2/E \ \cdots \ E_n/E] \quad (9)$$

式中， $E = (\sum_{i=1}^n |E_i|^2)^{\frac{1}{2}}$ 。

设铁路信号设备故障为轨道电路故障、信号机故障、道岔故障和微机联锁故障，同时包括正常状态，它们的特征向量值如表1所示。

表1 铁路信号设备状态的特征向量

Tab. 1 Characteristic vector of railway signal equipment status

E_0	E_1	E_2	E_3	E_4	状态
0.122	0.232	0.212	0.117	0.317	正常
0.074	0.114	0.397	0.171	0.244	轨道电路故障
0.223	0.253	0.137	0.234	0.153	信号机故障
0.32	0.284	0.203	0.137	0.056	道岔故障
0.098	0.199	0.165	0.176	0.362	微机联锁故障

1.3 铁路信号设备故障分类器建立以及参数优化

采用数据挖掘技术的支持向量机建立铁路信号设备故障分类器。对于铁路信号设备故障诊断数据样本，引入函数映射 $\phi(x)$ 到高维空间中，在该空间中进行决策，即

$$f(x, w) = w^T f(x) + b \quad (10)$$

定义误差函数，即

$$L_\varepsilon(y) = \begin{cases} 0 & |f(x) - y| \leq \varepsilon \\ |f(x) - y| - \varepsilon & |f(x) - y| > \varepsilon \end{cases} \quad (11)$$

为了计算式(10)中的 w 和 b 的值，建立如下目标函数

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\varepsilon_i - \varepsilon_i^*) \quad (12)$$

式(12)的约束条件为

$$\begin{cases} y_i - w^T \phi(x) - b \leq \varepsilon + \xi_i^* \\ -y_i + w^T \phi(x) + b \leq \varepsilon + \xi_i \\ \xi_i, \xi_i^* \geq 0, i = 1, 2, \dots, n \end{cases} \quad (13)$$

式中， C 为正则化参数； ξ_i 和 ξ_i^* 为松弛变量。

引入Lagrange乘子 a_i 和 a_i^* ，将式(14)转换为对偶问题，具体为

$$\begin{aligned} \max R(\alpha_i, \alpha_i^*) = & -\varepsilon \sum_{i=1}^n (\alpha_i + \alpha_i^*) + \sum_{i=1}^n (\alpha_i^* - \alpha_i) y_i - \\ & \frac{1}{2} \sum_{i,j} (\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j) K(x_i, x_j) \end{aligned} \quad (14)$$

w 计算公式为

$$w = \sum_{i=1}^n (\alpha_i - \alpha_i^*) \phi(x_i) \quad (15)$$

引入 $K(x_i, x)$ 为核函数，那么支持向量机建立铁路信号设备故障分类器决策函数可表示为

$$\begin{aligned} f(x) = & \sum_{i=1}^n (\alpha_i - \alpha_i^*) \phi(x_i) \phi(x) \\ = & \sum_{i=1}^n (\alpha_i - \alpha_i^*) K(x_i, x) + b \end{aligned} \quad (16)$$

选择高斯径向基核函数，具体为

$$K(x_i, x) = \exp\left(-\frac{\|x_i - x\|^2}{\sigma^2}\right) \quad (17)$$

式中， σ 为高斯宽度参数。

支持向量机建立铁路信号设备故障分类器性能受到参数 c 、 σ 影响很大，当前它们的值主要通过经验法确定，导致铁路信号设备故障分类器的建立时间长，而且故障诊断误差大，因此本文采用蝙蝠算法^[11]优化参数 c 、 σ ，建立更优的铁路信号设备故障分类器。

1.4 基于数据挖掘的铁路信号设备故障自动诊断步骤

1) 采集铁路信号设备的工作状态数据。

2) 采用小波变换对原始铁路信号设备的工作状态数据进行分解，对噪声小波系数置0，然后通过小波重构得到无噪的铁路信号设备的工作状态数据。

3) 采用经验模态分解法提取铁路信号设备故障诊断特征，对特征采用式(9)进行归一化操作。

4) 确定支持向量机参数 c 、 σ 的取值范围，并初始化蝙蝠算法自身参数。

5) 蝙蝠算法通过模拟蝙蝠找食物机制，

得到一条从巢到食物之间的最优路径，对最优路径进行解码，得到支持向量机参数 c 、 σ 的最优取值。

6) 根据 c 、 σ 的最优取值，支持向量机对铁路信号设备的工作状态和特征向量之间的关系进行拟合，建立铁路信号设备的工作状态分类器。

2 铁路信号设备故障诊断的性能对比

2.1 铁路信号设备故障诊断测试平台

为了分析数据挖掘的铁路信号设备的诊断效果，建立仿真测试平台，硬件平台为4核CPU，16 GB内存，软件采用Windows 2010和Matlab 2017，选择BP神经网络、传统支持向量机的铁路信号设备诊断方法进行对比测试，铁路信号设备工作状态的样本数量如表2所示。

表2 铁路信号设备状态的样本数量

Tab. 2 Sample number of railway signaling equipment status

状态	训练样本数量	测试样本数量
正常	200	50
轨道电路故障	20	10
信号机故障	40	20
道岔故障	10	5
微机联锁故障	10	5

2.2 结果与分析

三种方法的铁路信号设备诊断准确度对比如图1所示。对图1进行分析可知：

1) BP神经网络的铁路信号设备诊断准确度最低，这是因为BP神经网络的学习原理不一样，其要求训练样本规模比较大，而且铁路信号设备诊断训练样本数量少，因此无法对铁路信号设备诊断分类器进行充分训练，使得一些铁路信号设备诊断的误差比较大，铁路信号设备诊断结果不理想。

2) 支持向量机的铁路信号设备诊断结果

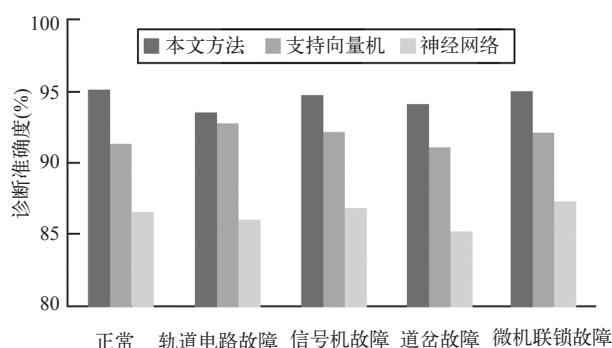


图1 铁路信号设备诊断准确度对比

Fig. 1 Comparison of diagnostic accuracy of railway signal equipment

要优于BP神经网络，这是因为支持向量机可以在小样本的条件下得到充分训练，但是由于参数采用随机方式确定，铁路信号设备诊断准确度要低于本文方法。由于本文方法引入了蝙蝠算法确定参数 c 、 σ 的最优取值，可以对铁路信号设备故障变化特点进行建模，构建铁路信号设备工作状态的分类器，得到了铁路信号设备故障诊断结果。

采用铁路信号设备故障诊断时间描述各种方法的工作效率，结果如图2所示。从图2可以看出，本文方法的优势十分明显，这是因为本文方法采用了蝙蝠算法和蝙蝠算法，减少了诊断时间。

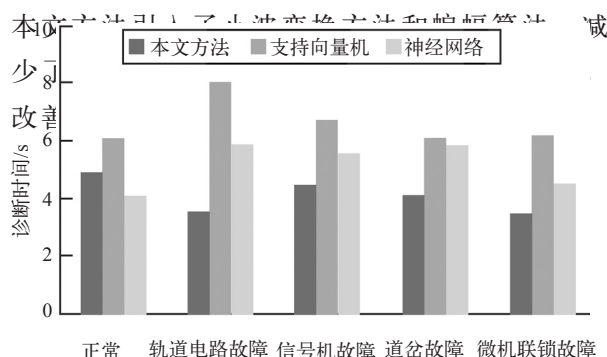


图2 铁路信号设备诊断时间对比

Fig. 2 Comparison of diagnosis time of railway signal equipment

3 结束语

铁路信号设备故障的类型多，故障难以识别，针对当前铁路信号设备故障诊断方法存在的弊端，提出了数据挖掘的铁路信号设备故障

自动诊断方法。在相同条件下,与经典铁路信号设备故障诊断方法相比,本文方法的铁路信号设备故障诊断准确度更高,速度更快,具有十分明显的优势。

参考文献

- [1] 宁滨,莫志松,李开成.高速铁路信号系统智能技术应用及发展[J].铁道学报,2019,41(3): 1-9.
- [2] 刘冬,魏彪,郑徐豪,等.铁路信号供电缺相故障的智能低压配电箱测控系统[J].测控技术,2016,35(4): 72-74.
- [3] 刘大为,郭进,王小敏,等.中国铁路信号系统智能监测技术[J].西南交通大学学报,2014,49(5): 904-912.
- [4] 赵文武,曾绍武,赵钦,等.BIM技术在铁路信号设备数据管理中的应用研究[J].铁道标准设计,2017,61(1): 127-133.
- [5] 张秉涛,王国荣.铁路信号设备质量诊断系统功能及运用[J].中国铁路,2008(7): 68-71.
- [6] 张喜,杜旭升,刘朝英.车站信号控制设备故障诊断专家系统的研究与实现[J].铁道学报,2009,31(3): 43-49.
- [7] 杨连报,李平,薛蕊,等.基于不平衡文本数据

挖掘的铁路信号设备故障智能分类[J].铁道学报,2018,40(2): 59-66.

- [8] 陆源,王瑞峰.信号设备故障诊断专家系统知识库的研究[J].微计算机信息,2007(28): 147-148.
- [9] 张霄,钱玉良,邱正,等.基于蜻蜓算法优化BP神经网络的燃气轮机故障诊断[J].热能动力工程,2019,34(3): 26-32.
- [10] 杨志淳,靖晓平,乐健,等.基于MI-PSO-BP算法的配电设备状态实时评估方法[J].电力自动化设备,2019,39(12): 108-113.
- [11] 郑云水,牛行通,康毅军.蝙蝠算法优化模糊神经网络的25 Hz相敏轨道电路故障诊断研究[J].铁道学报,2018,40(12): 93-100.

作者简介



张 硕

张硕(1985—),男,硕士,高级工程师,主要研究方向为铁道信号。

(收稿日期:2020-02-03)EA



典藏技术 革新应用

专业专注 时刻关注并推动中国电气技术的发展

《电气应用》是一本拥有光荣发展历程的全国优秀科技期刊。

自1982年创刊至今,已走过三十多年艰苦创业、改革发展的光辉历程。其所荣获的“全国优秀期刊、中国期刊方阵‘双奖期刊’和国家期刊奖‘百种重点期刊’、中文核心期刊”等诸多荣誉见证了编辑们的专业与勤奋,也博得了广大读者和业内专家的支持与厚爱,在众多电工类专业技术期刊中颇具影响力。



国内统一刊号: CN11-5249/TM 国际标准刊号: ISSN1672-9560 国内邮发代号: 82-341 国外发行代号: M1409