

基于概率模型算法的设备运行状态分析

褚红燕, 沈世斌

(南京师范大学 电气与自动化工程学院, 江苏 南京 210042)

[摘要] 从设备状态动态变化的本质出发, 采用概率神经网络方法构建设备状态概率模型, 描述设备从正常到故障的全过程. 通过对神经网络结构的分析和主要模型参数的计算, 以设备运行历史数据作为网络样本点, 在模型数据预处理的基础上建立了设备的概率模型, 该模型能较好地反映机械设备的运行状态. 通过现场数据的分析, 证明了该模型的合理性和正确性.

[关键词] 概率模型, 概率神经网络, 状态监测

[中图分类号] TP273+.4 [文献标识码] B [文章编号] 1672-1292(2005) 04-0034-04

Analysis of Equipment Running Condition Based on the Probabilistic Model Arithmetic

CHU Hongyan SHEN Shixin

(School of Electrical and Automation Engineering, Nanjing Normal University, Jiangsu Nanjing 210042, China)

Abstract Based on the essence of dynamic evolution of equipment condition, according to probabilistic neural networks (PNN), a probability model of equipment condition is adopted to describe the whole process of equipment evolution from normal to fault in the paper. By analyzing Neural Network construction, calculating the main model parameter and using the machine history data as the network sample, we establish the probability model of machine with the pretreatment model data as a basis. The machine equipment condition is well reflected by the probability model. The reasonability and correctness of the model are proved by analyzing the machine history data.

Key words probabilistic model, probabilistic neural networks, condition monitoring

设备状态分析的目的是保证设备在一定的环境和运行期间内安全、有效地实现其设计功能, 防止故障事故的发生, 并根据设备运行状态制定合适的维修计划, 提高企业运行的经济效益. 为实现上述目的, 就需要在设备运行过程中, 通过各种监测手段获取设备的工作状态信息, 然而由于现场运行环境、测量噪声和各种干扰的存在, 所采集和分析的特征量往往都是随机变量, 这些随机变量所蕴含的规律性无法用确定性的数学公式来表达, 只是具有统计意义上的描述.

概率密度函数给出了信号取不同幅值大小的概率, 因此设备某种监测指标的概率密度函数, 也就反映了该指标取不同值的概率. 设备状态发生变化时, 其特征量及相应地概率密度函数也必然作相应的改变, 相反, 特征量概率密度函数的改变, 也必

然反映了设备运行状态的改变.

确保机械设备的正常运转、防止重大事故的发生并延长其使用寿命早已引起了学术界和企业界的重视. 而随着设备监测诊断技术的不断发展, 对大型设备的在线监测技术和对中、小型机电设备的离线点检技术已逐步在企业中开始实施. 许多企业都已积累形成了大量的设备运行历史数据, 然而数据并不等于信息, 如何真正利用数据, 为企业设备管理、维修诊断提供切实有效的手段是目前急需解决的问题. 传统的设备运行状态分析主要采用静态指标来划分状态, 设定阈值. 设备状态的概念不清, 阈值适应性差, 漏报、谎报严重, 工程应用性差. 本文从设备状态动态变化的本质出发, 从设备运行的历史数据入手, 采用概率神经网络方法构建设备的状态概率模型, 为准确判断设备的运行状态提供了

有效的方法, 在实际应用中取得了较好的效果, 并为动态阈值设置和构建设备状态自适应报警线奠定了基础^[6]。

1 机械设备建模的基本假设

文献[5]中对机械设备的发展趋势做了分析, 得出以下结论: 机械故障的发生有其自身的规律性, 机械状态在大多数情况下是连续变化的, 即遵循正常 \Rightarrow 异常 \Rightarrow 故障的发展过程, 而且从正常到故障的中间演变过程是一种不稳定的状态, 发展非常迅速。对于以点检方式监测管理的设备来说, 由于是离线监测方式, 当设备经历着从正常 \Rightarrow 异常 \Rightarrow 故障状态的演变过程时, 所采集的历史数据必然包括大量正常状态下的数据、极少数异常状态下的数据和少量故障状态下的数据。当对这些历史数据作统计意义上的计算时, 会发现历史数据以很大的概率处于正常状态, 而以较小的概率处于故障状态, 在这两种状态之间包括一段概率分布极小的区域, 这个区域即是异常状态范围, 它是由设备经历异常状态的过程非常迅速而导致的。

因此, 由于故障模式的多样性、机械设备结构的复杂性、现场环境造成的随机性等因素的影响, 单纯地分析监测数据已很难准确地描述出设备的状态, 而模型化的思想则为解决这一问题提供了新的思路。从上面的分析可以得出结论: 即历史数据的概率分布与设备的运行状态有良好的对应关系, 这为用概率模型分析设备的状态提供了理论基础。

2 概率神经网络方法

建立设备状态概率模型的关键问题是寻找一种能准确地反映设备状态与历史数据概率分布之间关系的数学方法。概率神经网络由 Specht 于 1990 年提出, 是将基于 Bayes 理论的分类网络和概率统计思想组合起来而构成的, 这种网络具有快速优化能力、稳定的收敛性能和对现场数据的柔性追加等特点, 使得它在故障诊断、模式识别、非线性处理、趋势预报等领域有着广泛的应用。

概率神经网络的典型结构如图 1 所示, 它由输入层、模式层和求和层 3 层网络结构组成。在网络训练时, 网络直接存储训练样本向量作为网络的模式样本向量, 而不作任何修改, 只需对高斯函数的平滑因子进行经验和统计的估计, 过程极为简单。训练样本的选取原则是应包括设备大量正常状态下的数据、极少数异常状态下的数据和少量故障状态下的数据; 在网络工作时, 待识别样本 X 由输入

层直接送到模式层各个类别单元中, 在模式层中进行向量 X 和 x_i 的点积, 完成非线性处理后再送入求和层中; 求和层各单元与相应类别的模式单元相连, 并且依据 Parzen 方法求和估计各类的概率。

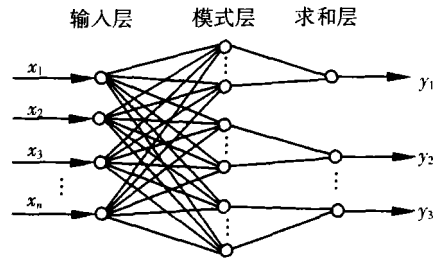


图 1 概率神经网络结构图

Parzen 方法是指 Parzen 于 1968 年提出的条件概率的计算方法, 其函数表达式为^[1]:

$$p\left(\frac{X}{f_k}\right) = \frac{1}{(2\pi)^{\frac{m}{2}} \sigma^m |f_k|} \sum_{x_i \in f_k} \exp\left[-\frac{\|X - x_i\|^2}{2\sigma^2}\right]$$

其中, X 是识别样本向量, $\{x_i\}$ 是模式样本向量, m 是向量维数, σ 是平滑因子, $|f_k|$ 是分类模式的数量。这种条件概率的估计方法表现为计算各高斯型节点函数的和。每个节点中心为训练样本模式点, 平滑因子就是高斯型函数的标准偏差 σ 。同时, 一个节点函数对条件概率的影响是随识别样本向量与模式样本向量间的距离平方呈指数变化的, 即节点函数对条件概率的影响具有局域性, 当一个节点间距离较大时, 该节点对条件概率的影响可忽略。

从 Parzen 方法中可以看出, 平滑因子的大小关系到整个概率密度分布函数的变化。图 2 所示为二维样本向量时平滑因子对概率模型的影响。在图中, 平滑因子 σ 由小变大, 所对应的概率密度分布函数具有明显的差异。由此可知, 优化平滑因子是构建设备状态概率模型的关键。

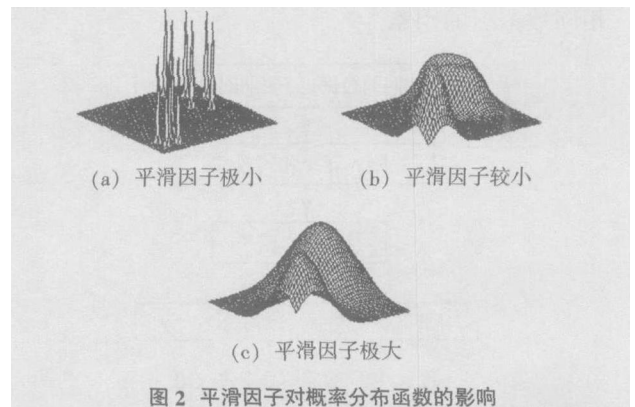


图 2 平滑因子对概率分布函数的影响

Cain 通过大量数据的实验处理, 表明平滑因子与样本间的平均最小距离有关, 而且被认为是一种既简单又行之有效的方法^[4], 因此本文中平滑因子估计采用其表达式:

$$\sigma = g \times \frac{1}{N} \sum_{i=1}^N m \ln d_j$$

其中, N 为样本层的数据个数; $d_j = |x_i - x_j|$ $i, j = 1, 2, \dots, N; i \neq j$ g 为经验系数.

3 模型构建的数据预处理

通过各种监测手段获得的设备运行数据并不能直接地用于构建概率模型, 必须进行一定的预处理, 使数据更易于构成概率神经网络中样本层的样本点. 数据预处理包括以下几个方面: 归一化处理, 相同数据的处理, 数据量偏少时的处理, 数据量过多时的处理. 一般情况下, 对前两个方面的处理总是必须进行的, 后两个方面则可根据实际情况分别进行. 本文只进行前两个方面的处理.

3.1 归一化处理

Specht在构造基本概率神经网络时规定对本向量必须进行归一化处理^[1], 归一化处理就是通过计算将历史数据的范围限定为 $[0, 1]$, 其处理过程为:

$$x_i = \frac{x_i - m \ln(x)}{m \max(x) - m \ln(x)} \quad i = 1, 2, \dots, N$$

其中, $\{x_i\}$ 为设备运行历史数据, $m \ln(x)$ 、 $m \max(x)$ 分别为样本向量的最小值与最大值, N 为数据个数, $\{x_i\}$ 为归一化后的数据.

3.2 相同数据的处理

由 Bayes理论可知, 我们在计算条件概率时都是假定其先验概率是等概率分布的. 而设备数据的相同点则说明同一个样本点出现了多次, 即其先验概率分布变大了. 因此, 在本文中对相同数据的处理过程如图 3 所示, 这样既避免了相同的数据点对平滑因子估计的影响, 又在建立概率模型中考虑了相同数据点的因素.

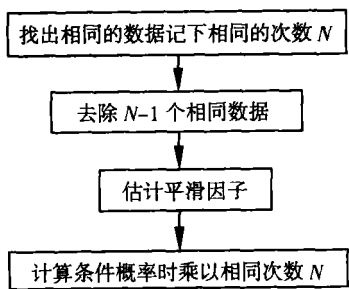


图 3 相同数据的处理流程图

4 模型的建立与分析

根据设备运行的历史数据, 可以把它们分为两类: 一类是设备投入使用后没有发生任何故障, 这

类设备的历史数据只包括正常运行范围内的数据, 这些数据大都集中在一个比较小的范围内, 因此此时概率神经网络样本层各样本点比较接近, 即样本点间的距离较小, 所计算出的平滑因子也较小, 每个样本点对整个概率模型的影响范围也较小; 另一类是设备曾发生一次或多次故障, 这类设备的历史数据包括了大量正常运行时的数据和少量故障状态下的数据, 而设备故障时, 状态的变化往往会导致数据远离正常范围, 反映在概率模型中, 则是网络样本层中除包括大量分布比较集中的样本点外, 还包括若干个比较分散的样本点, 估计出的平滑因子较大, 即每个样本点对整个概率模型曲线的影响范围较大.

图 4 是根据某厂减速机设备的振动速度测量值历史数据及其概率模型图, 其间共采集 77 组数据, 通过分析其测量值 (ms), 得出概率模型图. 该设备一直处于正常状态, 从概率模型图中可以看出, 正常设备的状态概率模型是在小范围内的概率分布函数. 它们的样本点比较集中, 平滑因子也较小, 即每个样本点对整个概率模型的影响较小, 但由于数据都大量集中在均值附近, 概率模型在历史数据平均值附近的概率达到最大. 只要设备处于正常运行状态, 则通过监测工具所采集的数据或通过波形数据计算出的指标有很大的概率发生在均值附近, 而远离均值的数据只是小概率事件, 因此, 此时的概率模型是完全符合设备运行的实际情况的.

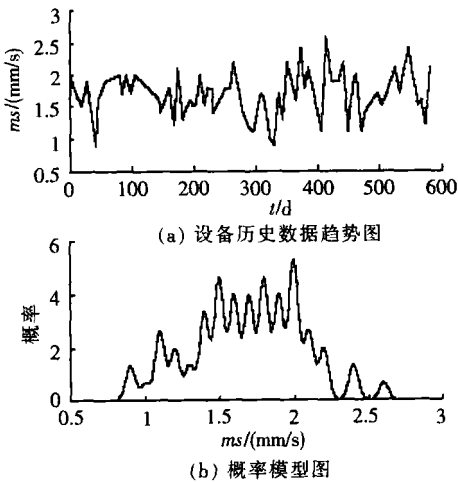


图 4 设备历史数据及概率模型图 1

图 5 是某厂减速机设备的振动速度测量值的设备历史数据及概率模型图, 共监测 78 组数据, 通过分析其测量值 (ms), 得出概率模型图. 该设备运行过程中发生了一次故障, 从概率模型图中可以看出, 由于故障数据偏离正常范围较远, 而且故障数据的分布范围较广, 即样本点间的距离较大, 估

计出的平滑因子往往都较大, 导致每个样本点的影响范围加大. 而且由于故障数据与正常数据间的距离较远, 它们之间的相互影响非常小, 导致在概率模型图中出现一个在正常范围和故障范围之间概率非常小的区域. 这个区域正是设备从正常状态到故障状态过渡的不稳定区域. 从设备的劣化过程来看, 设备出现在这种不稳定状态的时间较短, 因而在概率模型图中也是一个概率极小的范围, 此时的概率模型也是符合设备运行的实际情况的.

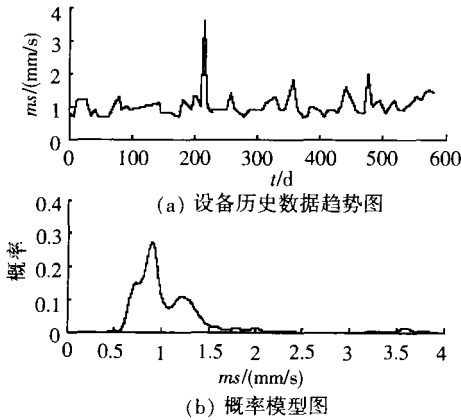


图 5 设备历史数据及概率模型图 2

图 4 和图 5 分别代表设备投入使用后没有发生任何故障和设备曾发生一次和多次故障的设备运行历史数据及所构建的概率模型. 从图中可以看出设备状态的概率模型客观地反映了设备的运行状态. 正常设备的概率模型与故障设备的概率模型之间存在着明显的区别, 因此完全可以将该概率模型作为对设备状态判定的有效判据.

5 结论

概率神经网络是以概率统计思想和贝叶斯优化规则而构成的分类神经网络, 本文介绍了概率神经网络的原理, 阐述了平滑因子的估计方法, 通过对数据的预处理和概率神经网络的建模, 分别对只包括正常状态数据以及包括正常和故障状态数据建立了概率模型, 并比较了它们的异同点. 对这两种数据所建立的概率模型既有相似的地方, 又有明显的区别, 但通过分析证明都符合现场设备运行的实际情况.

[参考文献]

- [1] Donald F Specht. Probabilistic neural networks and the polynomial adaline as complementary techniques for classification [J]. IEEE Transactions on Neural Networks 1990, 1(1): 111-121.
- [2] 徐光华, 屈梁生. 基于概率神经网络的机组状态多步预报方法 [J]. 西安交通大学学报: 自然科学版, 1999, 33(7): 85-87.
- [3] 叶志峰, 孙健国. 基于概率神经网络的发动机故障诊断 [J]. 航空学报, 2002, 23(2): 155-157.
- [4] Bibb Cain J. An improved probabilistic neural networks and its performance relative to other model [C] // Application of Artificial Neural Networks. SPIE Press, 1990.
- [5] 邱立鹏. 设备剩余寿命的预测与分析 [D]. 大连: 大连理工大学, 2000.
- [6] 高洪青. 基于概率模型的设备状态自适应评估与预测技术 [D]. 西安: 西安交通大学, 2003.

[责任编辑: 严海琳]