

可拓理论与SOM网络相结合的多故障诊断方法

文天柱, 许爱强, 陈育良

海军航空工程学院, 烟台 264001

摘要 针对采用SOM网络进行多故障诊断时,要求多故障模式相似且不包含标准故障输出的限制,提出将SOM网络与可拓理论相结合的多故障诊断方法。首先采用SOM网络对训练样本进行聚类,得到故障模式及其聚类中心。然后针对每种故障模式的每个特征构造在聚类中心处取得最大值的关联函数,并以各特征的关联函数值为基础,设计多故障评价指标实现多故障诊断。最后采用汽轮发电机组振动信号的频谱数据对算法进行验证,结果表明该方法能够正确识别待诊断样本的单故障和多故障模式,具有可行性。

关键词 多故障诊断;SOM网络;可拓理论;关联函数

中图分类号 TP206.3

文献标志码 A

doi 10.3981/j.issn.1000-7857.2014.34.008

Multi-fault Diagnosis Method Based on Combination of Extension Theory and SOM Network

WEN Tianzhu, XU Aiqiang, CHEN Yuliang

Naval Aeronautical and Astronautical University, Yantai 264001, China

Abstract This paper proposes a multi-fault diagnosis algorithm combining SOM network with extension theory to meet the requirement that multi-fault modes should be similar and do not contain the standard fault output when SOM network is used for multi-fault diagnosis. First, the training samples are clustered by SOM network, and the fault modes and these clustering centers can be obtained. Second, the dependent function of each feature for each fault mode is set up where the maximum value can be obtained at the clustering center. Next, the evaluation index of multi-fault modes is designed for multi-fault diagnosis, which is based on the dependent function values of features. Finally, the spectrum data of vibration signal of steam turbine generator unit is adopted to verify the algorithm. The results show that both single-fault mode and multi-fault modes can be correctly distinguished by this method, so the algorithm is feasible.

Keywords multi-fault diagnosis; SOM network; extension theory; dependent function

随着武器装备系统化、综合化、信息化程度的不断提高,对其进行维修、保障的难度也急剧增加。故障诊断技术作为降低武器装备维修保障费用、延长使用寿命、保持战备完好性的重要手段,得到了广泛关注。人工神经网络是一种基于知识的故障诊断方法,具有良好的非线性和并发性处理能力,以及联想和记忆功能,比较典型的神经网络模型有前馈神经网络模型、反馈神经网络模型、自组织神经网络模型和

随机神经网络模型等。自组织映射(self-organizing map, SOM)网络是一种典型的自组织神经网络模型,由于采用竞争式的无导师监督学习规则,所以在训练过程中不需要已知样本的故障类型,就可以自动对输入模式进行聚类,具有自组织和自适应能力^[1]。文献[2]应用自组织神经网络对汽轮机回热系统进行故障诊断;文献[3]将自组织神经网络用于轴流分机不对中故障诊断;文献[4]将自组织神经网络用于汽轮发

收稿日期:2014-09-25; 修回日期:2014-10-17

基金项目:“泰山学者”建设工程专项

作者简介:文天柱,博士研究生,研究方向为复杂军用电子装备自动测试技术和智能故障诊断技术,电子信箱:wentianzhu1987@aliyun.com

引用格式:文天柱,许爱强,陈育良.可拓理论与SOM网络相结合的多故障诊断方法[J].科技导报,2014,32(34):58-61.

动机组的多故障诊断。这些文献都是根据待诊断样本在SOM网络输出层的拓扑位置进行诊断,如果输出位置与标准故障输出重合,则诊断为与该标准故障相同类型的单故障;如果输出位置与多个标准故障输出相邻但不重合,则诊断为与相邻标准故障类型相同的多故障。而事实上,设备的多故障模式并不一定都是相似的,而且对于包含标准故障输出的多故障模式,该方法也不适用。

可拓学是专门用于解决矛盾和不相容问题的一门横断学科,它采用定性与定量相结合的方法对事物的发展变化情况进行描述^[9]。采用可拓学中的关联函数可以对待诊断样本属于各故障模式的程度进行定量计算,从而判断待诊断样本的故障类型,但关联函数构造困难,尤其是各故障模式的中心不易确定。为此本文将可拓理论与SOM网络相结合,首先用SOM网络对样本进行聚类,得到各种故障模式中心,然后构造最大值为各故障模式中心的关联函数,并设计基于关联函数值的故障评价指标,最后采用设计的故障评价指标对待诊断样本进行多故障诊断。

1 SOM网络聚类

1.1 SOM网络结构

SOM网络包括输入和输出两层,并在输出层引入网络拓扑结构模拟生物学中的侧抑制现象。输出层采用二维拓扑结构的SOM网络如图1所示。

图1所示SOM网络中,输入层的每个神经元通过权向量与任意输出神经元相连,输出层按行列排成矩阵形式,并通过权向量实现任意两个输出神经元之间的侧抑制。

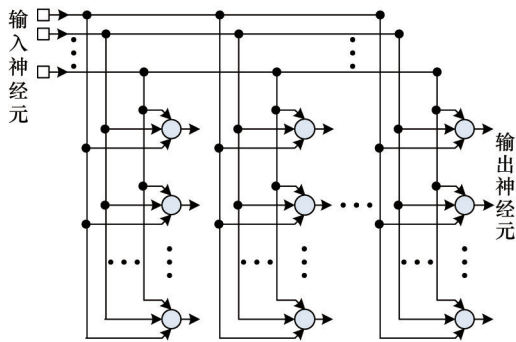


图1 SOM网络结构

Fig. 1 SOM network structure

1.2 聚类过程

1) 初始化:给定训练样本集 $\{S_1, S_2, \dots, S_{N_1}\}$, 样本集中任意样本 S_k 包含 m 个特征 c_1, c_2, \dots, c_m 与输入神经元对应, 样本 S_k 的特征值向量为 $[x_{k1}, x_{k2}, \dots, x_{km}]$ 。采用公式

$$x'_{ki} = \frac{x_{ki} - \min_{k=1,2,\dots,N_1} x_{ki}}{\max_{k=1,2,\dots,N_1} x_{ki} - \min_{k=1,2,\dots,N_1} x_{ki}}$$

对全部样本的特征值向量进行标准化,其中 x_{ki} 和 x'_{ki} 分别表示样本 S_k 特征 c_i 标准化前后的值, $k=1,2,\dots,N_1$,

$i=1,2,\dots,m$ 。采用随机值初始化SOM网络的权值矩阵 $W=[w_{ij}]_{m \times n}$, w_{ij} 表示输入神经元 in_i 和输出神经元 out_j 之间的权值。

2) 计算获胜神经元:样本 S_k 在输出层任意神经元 out_j 的输出 $y_j(k) = \sum_{i=1}^m w_{ij} x_{ki}$, 根据输出最大化的最优匹配准则, 样本 S_k 对应的获胜神经元为 out_p , $p = \arg \max_{j=1,2,\dots,n} y_j(k)$ 。

3) 阈值学习:为了防止随机初始化权值导致的神经元“饿死”,引入阈值学习机制,对于经常获胜的神经元需要更大的输出才能获胜,而对于“饿死”的神经元只需要较小的输出就能获胜,通过计算神经元的获胜次数与整个训练样本个数的比值判断神经元是否经常获胜,即输出神经元 out_j 的阈值 $\delta_j = f\left(\frac{N_p(j)}{N_1}\right)$, 其中 $N_p(j)$ 表示输出神经元 out_j 的获胜次数。

4) 权值更新:对获胜神经元邻域内的全部神经元的权值进行更新,更新公式为

$$w'_{ij} = w_{ij} + \eta(x_{ki} - w_{ij})$$

其中, $\eta = ab$, a 表示SOM网络的学习率, b 表示获胜神经元的邻域函数,且

$$b = \begin{cases} 1 & j=p \\ 0.5 & j \neq p \text{ 且 } j \in N_p(d) \\ 0 & j \notin N_p(d) \end{cases}, j=1,2,\dots,n$$

其中, $N_p(d)$ 表示所有距离获胜神经元小于 d 的输出神经元构成的邻域。

权值更新包括排序和收敛两个阶段,在排序阶段学习率从 a_1 下降到 a_2 , 邻域大小由 $N_p(d)$ 下降到 $N_p(1)$, 通过对权值向量的不断调整,使输出层体现输入样本的分布情况。收敛阶段学习率从 a_2 继续缓慢减小,邻域保持不变,实现拓扑结构的微调,使其分布更均匀^[6]。

5) 终止条件:当迭代达到指定次数 $step$, 算法终止。

算法结束后得到训练样本在输出层的拓扑结构,据此可确定训练样本对应的故障类型,每种故障类型包含的全部获胜神经元的权值向量中心即为聚类中心。

2 基于关联函数的诊断

2.1 构造关联函数

假设SOM网络聚类后得到的故障类型为 F_1, F_2, \dots, F_{N_2} , 故障类型 F_i 的聚类中心为 $[v_{i1}, v_{i2}, \dots, v_{im}]$, 对每种故障类型的每个特征构造最大值为聚类中心的关联函数^[7,8], 则故障 F_i 特征 c_i 的关联函数为

$$k_i(x_i) = \frac{\rho(x_i, v_{ii}, V_{0ii})}{D(x_i, V_{0ii}, V_i)}$$

其中 $\rho(x_i, v_{ii}, V_{0ii})$ 表示在特征值 $x_i = v_{ii}$ 时取得最大值的侧距, $D(x_i, V_{0ii}, V_i)$ 表示特征值 x_i 关于经典域 V_{0ii} 和节域 V_i 的位置。

故障 F_i 特征 c_i 的节域 $V_i = \langle v_i^L, v_i^U \rangle$, 其中下限 $v_i^L = \min_{k=1,2,\dots,N_1} x_{ki}$, 上限 $v_i^U = \max_{k=1,2,\dots,N_1} x_{ki}$; 经典域 $V_{0ii} = \langle v_{0ii}^L, v_{0ii}^U \rangle$, 其中

$v_{0ii}^L = \min_{S_i \in F_i} x_{ki}$, $v_{0ii}^U = \max_{S_i \in F_i} x_{ki}$, 则当 $v_{ii} < \frac{v_{0ii}^L + v_{0ii}^U}{2}$ 时, $\rho(x_i, v_{ii}, V_{0ii})$ 为左侧距

$$\rho(x_i, v_{ii}, V_{0ii}) = \begin{cases} v_{0ii}^L - x_i & x_i \leq v_{0ii}^L \\ \frac{v_{0ii}^U - v_{ii}}{v_{0ii}^L - v_{ii}}(x_i - v_{0ii}^L) & v_{0ii}^L < x_i < v_{ii} \\ x_i - v_{0ii}^U & x_i \geq v_{ii} \end{cases}$$

当 $v_{ii} > \frac{v_{0ii}^L + v_{0ii}^U}{2}$ 时, $\rho(x_i, v_{ii}, V_{0ii})$ 为右侧距

$$\rho(x_i, v_{ii}, V_{0ii}) = \begin{cases} v_{0ii}^L - x_i & x_i \leq v_{ii} \\ \frac{v_{0ii}^U - v_{ii}}{v_{0ii}^L - v_{ii}}(v_{0ii}^U - x_i) & v_{ii} < x_i < v_{0ii}^U \\ x_i - v_{0ii}^U & x_i \geq v_{0ii}^U \end{cases}$$

当 $v_{ii} = \frac{v_{0ii}^L + v_{0ii}^U}{2}$ 时, 侧距退化为最大值在经典域中点的距

$$\rho(x_i, V_{0ii}) = \left| x_i - \frac{v_{0ii}^L + v_{0ii}^U}{2} \right| - \frac{1}{2}(v_{0ii}^U - v_{0ii}^L)$$

同时, 对于特征值 x_i 关于经典域 V_{0ii} 和节域 V_i 的位置 $D(x_i, V_{0ii}, V_i)$, 当 $\rho(x_i, v_{ii}, V_{0ii}) \neq \rho(x_i, v_{ii}, V_i)$, $x_i \notin V_{0ii}$ 时

$$D(x_i, V_{0ii}, V_i) = \rho(x_i, v_{ii}, V_i) - \rho(x_i, v_{ii}, V_{0ii})$$

当 $\rho(x_i, v_{ii}, V_{0ii}) \neq \rho(x_i, v_{ii}, V_i)$, $x_i \in V_{0ii}$ 时

$$D(x_i, V_{0ii}, V_i) = \rho(x_i, v_{ii}, V_i) - \rho(x_i, v_{ii}, V_{0ii}) + v_{0ii}^L - v_{0ii}^U$$

当 $\rho(x_i, v_{ii}, V_{0ii}) = \rho(x_i, v_{ii}, V_i)$ 时

$$D(x_i, V_{0ii}, V_i) = v_{0ii}^L - v_{0ii}^U$$

2.2 计算故障评价指标

对于任意待诊断样本 Z_s , $s = 1, 2, \dots, N_3$, 其特征值向量为 $[x_{s1}, x_{s2}, \dots, x_{sm}]$, 则其关于故障类型 F_i 的评价指标为

$$K_{st} = \sum_{i=1}^m \lambda_i k_i(x_{si})$$

其中 λ_i 表示第 i 个特征的权重, 在没有先验知识的条件下 $\lambda_i = \frac{1}{m}$, $k_i(x_{si})$ 表示待诊断样本第 i 个特征关于故障类型 F_i 的关联函数。

当存在评价指标 $K_{st} \geq 0$, $t = 1, 2, \dots, N_2$ 时, 表示待诊断样本 Z_s 的故障类型为 F_t 。如果对于全部故障类型 F_i , 都有 $K_{st} < 0$, 则无法直接确定待诊断样本 Z_s 的故障类型, 原因可能是训练样本没有体现所有可能的故障类型或故障类型特征的经典域, 针对这种情况选择故障评价指标较大的故障类型作为诊断结果, 即如果 $\max_{t=1, 2, \dots, N_2} (K_{st}) - K_{st} < \delta$, 则待诊断样本 Z_s 的故障类型为 F_t , 一般取 $\delta = 0.01$ 。

3 实例验证

采用文献[9]中汽轮发电机组振动信号频谱数据进行实例验证。共有 18 组数据, 每组数据包含 8 个特征指标, 分别是 $0.01f \sim 0.39f$, $0.40f \sim 0.49f$, $0.50f$, $0.51f \sim 0.99f$, $1.00f$, $2.00f$, $3.00f \sim 5.00f$ 和 $> 5.00f$ (f 为汽轮发电机组的旋转频率)。采用 SOM 网络对 18 组样本进行聚类, 由于一般要求输出神经元个数要略大于样本个数, 所以设置输出神经元拓扑结构为 4 行 5

列共 20 个, 学习率 $a_1 = 0.9$, $a_2 = 0.02$, 获胜神经元初始邻域为 $N_p(3)$, 迭代 $step = 1000$ 次后得到训练样本在输出层的聚类结构如图 2 所示。

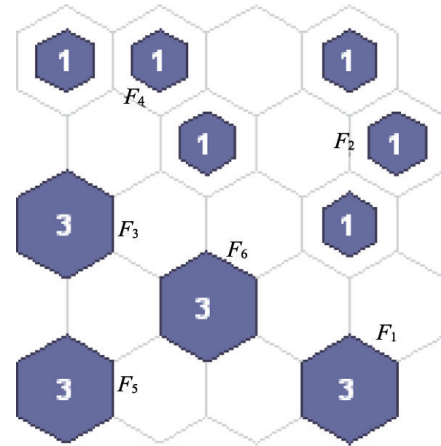


图 2 训练样本聚类结果

Fig. 2 Clustering result of training samples

图 2 中, 20 个输出神经元按从左到右, 从下到上的顺序依次标号, 获胜神经元标注的数字表示对应样本的数量, 最终确定 6 种故障类型, 每种故障类型包含 3 个故障样本, 第 1 号输出神经元表示故障类型 F_5 , 对应训练样本为 S_{13}, S_{14}, S_{15} ; 第 4 号输出神经元表示故障类型 F_1 , 对应训练样本为 S_1, S_2, S_3 ; 第 6 号输出神经元表示故障类型 F_6 , 对应训练样本为 S_{16}, S_{17}, S_{18} ; 第 9 号输出神经元表示故障类型 F_3 , 对应训练样本为 S_7, S_8, S_9 ; 第 12、16、20 号输出神经元表示故障类型 F_4 , 对应训练样本为 S_{10}, S_{11}, S_{12} ; 第 14、17、18 号输出神经元表示故障类型 F_2 , 对应训练样本为 S_4, S_5, S_6 。由原文献可知, 聚类结果与实际情况相符, 并且 F_1, F_2, \dots, F_6 表示的故障类型分别是不对中, 不平衡, 轴承座松动, 转子碰摩, 油膜涡动和蒸汽涡动。各种故障聚类中心见表 1。

用得到的各故障聚类中心构建关联函数, 对原文献中的 3 组待诊断数据进行诊断, 得到 3 组待诊断数据对应 6 种故障类型的故障评价指标, 见表 2。

由表 2 可知, 由于测试样本针对每种故障类型的评价指标 $K_{st} < 0$, $s = 1, 2, 3$, $t = 1, 2, \dots, 6$, 所以计算 $\max_{t=1, 2, \dots, N_2} K_{st} - K_{st}$ 作为待诊断样本的故障评价指标, 见表 3。

取 $\delta = 0.01$, 由表 3 可知, 待诊断样本 Z_1 的故障类型是 F_1 , 待诊断样本 Z_2 的故障类型是 F_2 , 待诊断样本 Z_3 的故障类型是 F_1 和 F_2 , 诊断结果与文献[9]中的一致, 说明采用本文提出的方法进行多故障诊断是可行的, 但由于多故障样本数量少, 无法对方法进一步验证。另外, 由于 SOM 网络的初始权重是随机数, 所以每次聚类后输出神经元的拓扑结构不同, 但训练样本的聚类结果相同, 所以对最后的多故障诊断结果没有影响。

表1 各故障模式聚类中心

Table 1 Clustering centers of each fault mode

特征	F_1	F_2	F_3	F_4	F_5	F_6
c_1	0.015	0.017	0.900	0.107	0.107	0.023
c_2	0.010	0.017	0.020	0.023	0.840	0.343
c_3	0.015	0.003	0.020	0.040	0.007	0.107
c_4	0.020	0.007	0.003	0.137	0.113	0.643
c_5	0.405	0.800	0.043	0.313	0.033	0.010
c_6	0.565	0.023	0.030	0.110	0.017	0.003
c_7	0.100	0.067	0.010	0.113	0.027	0.010
c_8	0.020	0.010	0.150	0.247	0.013	0.010

表2 待诊样本故障评价指标

Table 2 Evaluation index of samples for diagnosis

待诊样本	F_1	F_2	F_3	F_4	F_5	F_6
Z_1	-0.357	-0.395	-0.370	-0.507	-0.488	-0.628
Z_2	-0.221	-0.038	-0.210	-0.296	-0.230	-0.363
Z_3	-0.334	-0.326	-0.367	-0.449	-0.493	-0.626

表3 待诊样本故障评价指标

Table 3 Evaluation index of samples for diagnosis

待诊样本	F_1	F_2	F_3	F_4	F_5	F_6
Z_1	0	0.038	0.013	0.150	0.131	0.271
Z_2	0.183	0	0.172	0.258	0.192	0.325
Z_3	0.008	0	0.041	0.123	0.167	0.300

4 结论

提出了一种将可拓理论与SOM网络相结合的多故障诊断方法,首先将训练样本输入到SOM网络进行聚类,然后针对聚类后得到的故障模式和每种故障模式中心,构建在聚类中心处取得最大值的关联函数,并利用每种属性的关联函数值设计多故障诊断的评价指标,最后采用汽轮发电机组振动信号频谱数据进行实例验证,结果表明该方法能够正确识别待诊断样本的单故障和多故障类型,但由于训练样本和待诊断样本都不充足,所以还需对方法进一步验证。

参考文献(References)

- [1] Kohonen T. Essentials of the self-organizing map[J]. Neural Networks, 2013, 37: 52-65.
- [2] 吴俊芬, 胡念苏, 赵瑜. 基于自组织特征映射神经网络的汽轮机回热系统故障诊断[J]. 武汉大学学报: 工学版, 2003, 36(4): 99-102.
Wu Junfen, Hu Niansu, Zhao Yu. Fault diagnosis of steam turbine regenerative system based on self-organizing neural network[J]. Engineering Journal of Wuhan University, 2003, 36(4): 99-102.
- [3] 尤丽静, 陈在平. 基于自组织神经网络的轴流风机不对中故障诊断研究[J]. 天津理工大学学报, 2010, 26(1): 68-70.

You Lijing, Chen Zaiping. Misalignment fault diagnosis of the axial flow fan based on self-organizing neural network[J]. Journal of Tianjin University of Technology, 2010, 26(1): 68-70.

- [4] 张彼德. 汽轮发电机组振动多故障诊断模型及方法研究[D]. 重庆: 重庆大学, 2002: 36-38.
Zhang Bide. Study on diagnosis models and methods of steam turbine generators-Set's vibration multiple fault[D]. Chongqing: Chongqing University, 2002: 36-38.
- [5] 杨春燕, 蔡文. 可拓工程[M]. 北京: 科学出版社, 2007.
Yang Chunyan, Cai Wen. Extension engineering[M]. Beijing: Science Press, 2007.
- [6] 邹云峰, 吴为麟, 李智勇. 基于自组织映射神经网络的低压故障电弧聚类分析[J]. 仪器仪表学报, 2010, 31(3): 571-576.
Zou Yunfeng, Wu Weilin, Li Zhiyong. Cluster analysis of arc fault in low-voltage based on SOM neural network[J]. Chinese Journal of Scientific Instrument, 2010, 31(3): 571-576.
- [7] 王永林. 关于可拓评价中几种关联函数的分析[J]. 计算机工程与应用, 2010, 46(12): 56-59.
Wang Yonglin. Analysis on some dependent functions of extension evaluation [J]. Computer Engineering and Applications, 2010, 46(12): 56-59.
- [8] 徐慧, 周建美, 蒋峥峥. 基于侧距的关联函数构造及应用[J]. 南通大学学报: 自然科学版, 2010, 9(2): 9-13.
Xu Hui, Zhou Jianmei, Jiang Zhengzheng. Construction and application of dependent function based on lateral distance[J]. Journal of Nantong University: Natural Science Edition, 2010, 9(2): 9-13.
- [9] 张彼德, 欧健, 孙才新, 等. 汽轮发电机多故障诊断的SOM神经网络方法[J]. 重庆大学学报: 自然科学版, 2005, 28(2): 36-38.
Zhang Bide, Ou Jian, Sun Caixin, et al. Application of SOM neural network in multiple fault diagnosis of turbo generator set[J]. Journal of Chongqing University: Natural Science Edition, 2005, 28(2): 36-38.

(责任编辑 刘志远)