

两种神经网络在变压器油色谱故障诊断中的应用

周亚明¹, 田野¹, 陈斌², 董书勇³

(1 宿迁供电公司, 江苏 宿迁 223800; 2 绍兴电力局修试工区, 浙江 绍兴 312000;

3 萧山供电局, 浙江 萧山 311201)

摘要:分析了 Levenberg-Marquardt(L-M)算法和 RPROP 算法的原理, 并将它们应用于充油电气设备故障诊断神经网络的训练。构造了满足要求的神经网络, 分别使用了 L-M 算法和 RPROP 算法进行训练和诊断。结果表明, 两种神经网络均能快速收敛到较高的精度; 对训练所得网络分别使用没有训练过的样本进行诊断, 两种算法的诊断准确率分别是 86.47%、92.00%。

关键词:变压器; 故障诊断; 神经网络

Abstract: The principles of the two algorithms Levenberg-Marquardt (L-M) algorithm and RPROP algorithm are introduced. Then they are used to diagnose the fault of transformer based on DGA. The suitable ANN structure is constructed and the above-mentioned two algorithms are used to train the ANN. The result indicates that both of them have the high accuracy and the fast convergence speed. The trained ANN is used to diagnose the fault of transformer and the accuracies of these two algorithms are 86.47% and 92.00%, respectively.

Key words: transformer; fault diagnosis; ANN

中图分类号: TM855 **文献标识码:** A **文章编号:** 1003-6954(2009)01-0084-04

大型电力变压器是电力系统中重要的输变电设备, 在电力系统中处于枢纽地位, 电力变压器的故障不仅影响电力系统的输电能力, 还可能造成电力系统的大规模停电, 给电力系统和国民经济带来巨大损失^[1]。因此准确诊断变压器故障具有较大的理论意义和实用价值。

目前电力系统中的大型变压器大多是以精炼矿物油作为绝缘和冷却的介质, 不同类型的故障及不同严重程度故障产生气体的类型和浓度是不同的, 其中一些气体能反映变压器故障的情况, 通常称这些气体为特征气体, 它们是 H₂、CH₄、C₂H₆、C₂H₄、C₂H₂、CO、CO₂。变压器油中溶解气体分析 (Dissolved Gas in-oil Analysis 简称 DGA)^[2~3]就是根据气相色谱方法获得油中各种特征气体的浓度, 然后根据获得的各种气体浓度判断故障的一种方法。由于油中溶解气体分析方法对于判断慢性局部潜在性缺陷十分有效, 而且该方法无需停电试验, 方便用于在线监测, 因此, 在 1997 年颁布执行的电力设备预试规程中, 把变压器油中溶解气体分析放到了首要的位置。

由于变压器油中溶解气体的形成涉及复杂的机理, 而且与电、热、机械等方面的作用相互耦合, 通过理论方法无法完全精确地确定从变压器油中溶解的

特征气体浓度到变压器所属的故障类型及严重程度之间的映射关系。多层前馈神经网络^[4~5]具有形成非线性映射、自学习、自适应等优点, 非常适合于逼近故障特征气体浓度到故障类型的映射, 所以它在基于油中溶解气体分析的变压器故障诊断中有着十分广泛的应用。下面使用了 Levenberg-Marquardt 算法^[6,7]和 RPROP 算法^[8,9]训练的多层前馈神经网络对变压器进行故障诊断, 构造了网络模型, 使用例子验证了两种算法的有效性。

1 L-M 算法

多层前馈神经网络出现的一种比较有效的算法是 BP 算法, 但 BP 算法从数学原理上是基于一阶偏导数, 存在着以下两个问题: ①学习步长选择困难。过小的学习步长导致收敛速度慢, 过大的学习步长导致振荡, 很难选择一个在全局范围内都合适的最优步长。②容易陷入局部极小点。网络误差曲面上存在许多局部极小点, 这些点上网络误差最小误差较大, 但权值和阈值对误差的偏导数却很小, BP 算法权值和阈值的调整量是它们对误差的一阶偏导数与学

习步长的乘积,所以在局部极小点上权值和阈值的调整量很小,导致网络很难跳出局部极小点。虽然人们采用变学习速率和加动量项进行了改进,但以上两个问题在一定程度上仍存在。

为此,引入了具有二阶收敛速度的 L-M 算法,该算法是一种专门用于误差平方和最小化的神经网络训练算法。设样本数目为 M、输出神经元数目为 L 则网络误差的平方和

$$E = \frac{1}{2} \sum_{i=1}^M e_i = \frac{1}{2} \sum_{i=1}^M \sum_{j=1}^L \epsilon_{ij}^2 \quad (1)$$

式中, e_i 为第 i 个样本输出误差的平方和; ϵ_{ij}^2 为第 i 个样本在第 j 个输出层神经元的误差。

设 W 为网络的可变参数组成的列向量,网络可变参数为 N 个,则其形状为 $N \times 1$,它也是函数 E 的自变量,对函数 E 进行泰勒展开式取二阶近似并考虑平稳点,则有

$$W(k+1) = W(k) - H(k)^{-1} \frac{\partial E}{\partial W}(k) \quad (2)$$

式中, H 称为 Hessian 矩阵, $H_{ij} = \frac{\partial^2 E}{\partial w_i \partial w_j}$, 其形状为 $N \times N$; 如果直接计算 Hessian 矩阵则称为牛顿法。但 Hessian 矩阵的直接计算及求逆都很麻烦,而且常常是病态的,所以 L-M 算法用下式对 Hessian 矩阵进行近似:

$$H = J^T J + \lambda I \quad (3)$$

式中, J 称为 Jacobian 矩阵,其元素 $J_{ij} = \frac{\partial \epsilon_i}{\partial w_j}$, 其形状为 $ML \times N$; λ 为引入的可变参数。

网络可变参数的调整公式为:

$$W = W - (J^T J + \lambda I)^{-1} J^T \epsilon(W) \quad (4)$$

式中, ϵ 为由误差组成的矩阵,其形状为 $ML \times 1$ 。

该算法根据误差 E 的变化情况自动调整 λ , 当 λ 很小时,变为牛顿法;当 λ 很大时则变为 BP 算法,此时步长为 λ^{-1} 。

L-M 算法根据网络误差的变化情况在基于一阶偏导数的标准 BP 算法和基于二阶偏导数的牛顿法之间自适应变化,它的收敛速度比 BP 算法快,而且网络训练陷入局部极小点的可能性小。且式 (3) 中的 Jacobian 矩阵 J 基于一阶偏导,即上式仅有一阶计算量,故该算法的计算量相对 Newton 算法要小一些。

2 RPROP 算法原理

除了基于二阶改进算法的 L-M 算法可以用于多层前馈神经网络的训练,基于启发式改进的 RPROP 算法也能提高神经网络的性能。

多层前馈神经网络的误差函数是一个以网络可变参数为自变量、网络误差为因变量的非线性多元函数,它是多维空间中的一个形状极为复杂的曲面,网络的误差曲面随可变参数的变化十分剧烈,有些区域非常平滑,应该选择较大的学习步长;有些区域非常陡峭,应该选择较小的学习步长^[5]。但标准 BP 算法在训练中学习步长保持恒定,导致了学习步长选择困难;学习步长过大则网络容易振荡;学习步长过小则网络收敛速度太慢。而且网络误差曲面上存在着很多局部极小点,在局部极小点上网络误差较大,但可变参数对误差的一阶偏导数接近于 0 由于 BP 算法网络可变参数的调整量是可变参数对网络误差的一阶偏导数与学习步长的乘积,而学习步长是某一较小的常数,它们的乘积仍然是一个较小的值。因此,在局部极小点上网络可变参数的调整量很小,这使网络经过多次训练后仍不能跳出局部极小点,训练不收敛。

设 n 为迭代次数, $\Delta_{ij}^{(n)}$ 为网络可变参数调整量的幅值, $\frac{\partial E^{(n)}}{\partial w_{ij}}$ 为网络可变参数对网络误差的一阶偏导数, $\Delta w_{ij}^{(n)}$ 为网络可变参数的调整量。 Δ_{ij} 的调整公式如下:

$$\Delta_{ij}^{(n)} = \begin{cases} \eta^+ * \Delta_{ij}^{(n-1)}, & \text{当 } \frac{\partial E^{(n-1)}}{\partial w_{ij}} * \frac{\partial E^{(n)}}{\partial w_{ij}} > 0 \\ \eta^- * \Delta_{ij}^{(n-1)}, & \text{当 } \frac{\partial E^{(n-1)}}{\partial w_{ij}} * \frac{\partial E^{(n)}}{\partial w_{ij}} < 0 \\ \Delta_{ij}^{(n-1)}, & \text{否则} \end{cases} \quad (5)$$

式中, η^+ 通常取 1.2; η^- 通常取 0.5。在计算出 Δ_{ij} 后,可用式 (6) 计算出可变参数的调整量:

$$\Delta w_{ij}^{(n)} = \begin{cases} -\Delta_{ij}^{(n)}, & \text{当 } \frac{\partial E^{(n)}}{\partial w_{ij}} > 0 \text{ 且 } \frac{\partial E^{(n)}}{\partial w_{ij}} * \frac{\partial E^{(n-1)}}{\partial w_{ij}} > 0 \\ +\Delta_{ij}^{(n)}, & \text{当 } \frac{\partial E^{(n)}}{\partial w_{ij}} < 0 \text{ 且 } \frac{\partial E^{(n)}}{\partial w_{ij}} * \frac{\partial E^{(n-1)}}{\partial w_{ij}} > 0 \\ -\Delta w_{ij}^{(n-1)}, & \text{当 } \frac{\partial E^{(n)}}{\partial w_{ij}} * \frac{\partial E^{(n-1)}}{\partial w_{ij}} < 0 \\ 0, & \text{否则} \end{cases} \quad (6)$$

网络的可变参数用式 (7) 调整:

$$w_{ij}^{(n+1)} = w_{ij}^{(n)} + \Delta w_{ij}^{(n)} \quad (7)$$

BP 算法的 $\Delta w_{ij}^{(n)}$ 等于 $\frac{\partial E^{(n)}}{\partial w_{ij}}$ 与学习步长的乘积,

而 RPROP 算法的 $\Delta w_{ij}^{(n)}$ 是由 $\Delta_{ij}^{(n)}$ 与 $\frac{\partial E^{(n)}}{\partial w_{ij}}$ 和 $\frac{\partial E^{(n-1)}}{\partial w_{ij}}$ 的符号决定, 与 $\frac{\partial E^{(n)}}{\partial w_{ij}}$ 和 $\frac{\partial E^{(n-1)}}{\partial w_{ij}}$ 的幅值无关, RPROP 算法与前者主要区别在于后者。当网络处于局部极小点或误差曲面较平的区域时 (即 $\frac{\partial E}{\partial w_{ij}}$ 的幅值很小), Δw_{ij} 不受 $\frac{\partial E}{\partial w_{ij}}$ 的影响仍然保持相对较大的值, 所以网络更有可能跳出局部极小点而收敛于全局极小点。同时 $\frac{\partial E}{\partial w_{ij}}$ 仅表示 w_{ij} 按照 $\frac{\partial E}{\partial w_{ij}}$ 方向使用足够小的步长时网络误差有最快的下降速度, $\frac{\partial E}{\partial w_{ij}}$ 的幅值与最优调整步长并没有必然的联系。RPROP 算法使用 $\frac{\partial E}{\partial w_{ij}}$ 的符号信息, 而没有使用 $\frac{\partial E}{\partial w_{ij}}$ 的幅值信息, 能根据网络训练情况自适应调整 Δw_{ij} , 避免了次要信息的干扰, 加快了网络的收敛。

与 BP 算法需要确定网络的学习步长不同, RPROP 算法不需要用户确定学习步长。 η^+ 和 η^- 选择缺省值即可, 需要用户确定的只有 $\Delta_{ij}^{(0)}$, 而且网络训练对于 $\Delta_{ij}^{(0)}$ 不敏感, 将它初始化为任何大于零的随机数网络也能快速地调整到最优值^[8], 通常将 $\Delta_{ij}^{(0)}$ 全部初始化为一个较小的数, 比如 0.1。网络参数的高度自适应避免了用户通过不断试验取最优值的过程, 是 RPROP 算法相对于 BP 算法的优点之一。

RPROP 算法的计算步骤如下。

- ① 令网络迭代次数 $n=0$, $\Delta_{ij}^{(0)}=0.1$, $\eta^+=1.2$, $\eta^-=0.5$, 确定训练目标 E_g 、最多可能迭代次数 N_{max} 。
- ② 计算网络可变参数对网络误差的一阶偏导数 $\frac{\partial E^{(n)}}{\partial w_{ij}}$, 根据式 (5) 计算 $\Delta_{ij}^{(n)}$ 。
- ③ 根据式 (6) 计算网络可变参数的调整量, 根据式 (4) 对网络可变参数进行调整。
- ④ 计算所有样本的误差平方和 E_n , $n=n+1$, 如果 $E < E_g$ 或 $n > N_{max}$ 则训练结束; 否则, 跳至 ②。

3 神经网络的实现及训练和诊断效果

3.1 网络结构

3.1.1 输入层

输入信息是氢气、甲烷、乙烷、乙烯、乙炔 5 种气体的浓度, 用 5 个神经元表示。输入的形式有直接输入各种气体浓度和对样本进行归一化两种方式, 前者容易导致样本空间过大, 进而导致网络的规模过大, 影响网络的正常训练和诊断; 文中选用后一种方式, 将上述 5 种气体浓度与气体浓度总和的比值作为网络的输入。

3.1.2 隐层

这里使用通常采用的单隐层前馈神经网络。网络隐层神经元数目对网络的学习能力和泛化能力有着很大的影响, 考虑到神经网络的样本数目是 50 个, 综合考虑了学习能力和泛化能力, 通过试验确定隐层使用 12 个神经元。

3.1.3 输出层

有 6 个神经元, 分别表示网络能诊断的 6 种故障类型, 即低温过热、中温过热、高温过热、局部放电、低能放电和高能放电。输出为 1 表示属于此类故障, 0 则表示不属于此类故障。

3.2 网络训练和诊断

为了比较 L-M 算法和 RPROP 算法的效果, 使用 50 个变压器油中溶解气体分析样本, 使用各种算法训练或构建网络 10 次, 所得训练误差与所需时间的平均值如表 1 所示。

表 1 多种算法的训练效果

	L-M 算法	RPROP 算法
训练误差平方和	5.26	5.00
所用时间平均值 (s)	28.31	42.57

从表 1 中可以看出 L-M 算法和 RPROP 算法训练所得误差平方和比较接近, 训练所需时间也比较接近。

使用基于 L-M 算法、RPROP 算法的神经网络对 150 组变压器油中溶解气体数据进行诊断 (使用 10 个训练所得网络诊断结果和作为最终结果), 准确率分别为 86.47%、92.00%, 选取 3 个典型的故障变压器油中溶解气体数据的诊断结果进行说明, 其相应的油中溶解气体含量如表 2 所示。

表 2 故障变压器油中溶解气体含量 ($\mu\text{L/L}$)

	H ₂	CH ₄	C ₂ H ₆	C ₂ H ₄	C ₂ H ₂
S ₁	195.7	58	16.4	91.6	96.9
S ₂	242.8	76.29	44.44	173.06	0
S ₃	408.8	56.1	4.6	93.6	282.6

表中: S₁、S₂、S₃ 分别表示样本 1、2、3。

神经网络诊断是采用 10 个神经网络训练而得, 将 10 次诊断结果输出的平均值计算, 得到故障情况如表 3、4 所示。使用两种神经网络对这 3 台变压器诊断结果分别为高能放电、高温过热和高能放电, 这与这 3 台变压器的吊罩检查结果是一致的, 表明了神经网络诊断结果的准确性。但比较而言, RPROP 算法训练的神经网络的诊断精度要稍高于 L-M 算法训练的神经网络的诊断结果。结论如下:

1 号变压器: 吊罩检查故障原因为变压器绕组第 10 匝间换位处, 因短路而使 A 相低压线圈局部严重烧伤, 内层大面积烧黑, 绝缘破损, 铁心片上有局部变黑现象。

2 号变压器: 吊罩检查故障原因为 A、B 相分接开关引线头过热, 包缠的绝缘纸被烧焦, 分接开关接触不良。

3 号变压器: 吊罩检查故障原因为选择开关静触头支撑板条变形, 触头接触不良, 造成严重的拉弧放电; 切换开关 A 相保护间隙击穿或短路, 环流烧断过渡电阻。

表 3 L-M 神经网络诊断所得结果

变压器	低温 过热	中温 过热	高温 过热	局部 放电	低能 放电	高能 放电
1 号	0.05	0.08	0.04	0.12	0.04	0.85
2 号	0.04	0.06	0.83	0.04	0.20	0.04
3 号	0.04	0.06	0.05	0.04	0.21	0.70

表 4 RPROP 神经网络诊断所得结果

变压器	低温 过热	中温 过热	高温 过热	局部 放电	低能 放电	高能 放电
1 号	0.06	0.07	0.04	0.04	0.05	0.90
2 号	0.04	0.06	0.89	0.06	0.05	0.04
3 号	0.06	0.08	0.80	0.12	0.09	0.08

4 结 论

使用 L-M 算法训练的 L-M 神经网络和使用 RPROP 算法训练的 RPROP 神经网络经过训练能有效获得色谱数据中隐藏的故障诊断规则, 训练精度能满足要求。对训练所得神经网络对未训练过实例进行的诊断表明, 神经网络能有效诊断变压器故障。但从训练速度上看 L-M 算法更快一些, 从诊断结果上看 RPROP 算法准确性更高一些。

参考文献

- [1] 王昌长. 电力设备的在线监测与故障诊断. 北京: 清华大学, 1996.
- [2] 操敦奎. 变压器油中气体分析与诊断 [M]. 武汉: 中国水利电力企业管理协会, 1987.
- [3] 孙才新, 陈根伟, 李俭, 等. 电气设备油中气体在线监测与故障诊断技术 [M]. 北京: 科学出版社, 2003.
- [4] 王胜财, 孙才新, 廖瑞金. 变压器色谱监测中的 BPNN 故障诊断法 [J]. 中国电机工程学报, 1997, 17(5): 322-325.
- [5] 黄鞠铭, 朱子述, 胡文华, 等. BP 网络在基于 DGA 变压器故障诊断中的应用 [J]. 高电压技术, 1996, 22(2): 21-23.
- [6] Hagan M T, Menhaj M. Training feedforward networks with the Marquardt algorithm. IEEE Transactions on Neural Networks 1994, 5(6): 989-993.
- [7] 赵登福, 林 谋, 李彦明, 等. 用 L-M 算法的神经网络诊断充油设备绝缘故障 [J]. 高电压技术, 2004, 30(7): 4-6, 16.
- [8] M Riedmiller, H Braun. A direct adaptive method for faster backpropagation learning: The RPROP algorithm. IEEE International Conference on Neural Networks San Francisco CA, USA, 1993, 586-591.
- [9] 章剑光, 周浩, 盛晔. 基于 RPROP 神经网络算法的主变 DGA 故障诊断模型 [J]. 电力系统自动化, 2004, 28(14): 63-66.

(收稿日期: 2008-10-14)

节约能源 保护环境