

模拟退火优化 SVM 参数的变压器故障诊断

谢荣斌¹, 张霖¹, 鄢小虎², 杨俊², 卢文华²

(1. 贵阳供电局, 贵阳 550001;

2. 国网电力科学研究院 武汉南瑞有限责任公司, 武汉 430074)

摘要: 利用群智能算法优化支持向量机 (SVM) 参数往往需要引入额外的变量, 使变压器故障诊断问题更加复杂, 对此提出一种自适应的模拟退火算法优化 (ASA) 支持向量机参数; 通过设计自适应的冷却进度表, 使得寻参的过程仅仅依赖于退火速率以及网格搜索粒度, 保持了较少的参数设置; 在相关数据集上的实验表明, 与已提出的粒子群算法 (PSO)、遗传算法 (GA) 相比, ASA 算法具有更快的收敛速度以及较好的诊断精度; 利用自适应的模拟退火算法能够较好的优化 SVM 参数并提高变压器故障诊断的精度。

关键词: 支持向量机; 模拟退火; 变压器故障诊断

Optimized Parameters by Simulated Annealing of SVM in Power Transformer Diagnosis

Xie Rongbin¹, Zhang Lin¹, Yan Xiaohu², Yang Jun², Lu Wenhua²

(1. Guiyang Power Supply Bureau, Guiyang 550001, China;

2. Wuhan Nanrui Company Limited, State Grid Electric Power Research Institute, Wuhan 430074, China)

Abstract: Optimizing the parameters of Supported Vector Machine generally introduces extra variables, causing the original problem more complicated. Hence, a Self-adaptive Simulated Annealing Algorithm is proposed. By means of designing a self-adaptive annealing schedules, the course of searching the best parameters only depends on two variables, annealing rate and grid search granularity, keeping relatively fewer parameters setting. As with the experiment shown on related datasets, ASA algorithm acquires the faster speed of convergence and decent diagnosis accuracy. Utilizing ASA acquires better parameters of SVM, while improving the accuracy of transformer diagnosis.

Keywords: supported vector machine; simulated annealing; power transformer diagnosis

0 引言

变压器是电力系统中极其重要的枢纽设备, 油中溶解气体分析 (dissolved gas analysis, DGA) 技术是变压器故障诊断的重要依据。基于 DGA 数据形成的传统方法在实际应用中往往误诊率很高。因此, 许多依赖于计算智能的方法被提出, 例如模糊理论^[1]、专家系统^[2]、灰色理论^[3]、支持向量机^[4]。

SVM 方法能够根据油中各种气体含量较精确的预测故障的类型。董明等人^[5]最早将支持向量机用于变压器故障诊断, 提出了分层决策电力变压器故障诊断模型。由于 SVM 的准确率严重依赖于参数的选择, 大量的研究集中在了如何寻找最优参数的问题上。基于遗传算法, 粒子群算法, 人工免疫算法等智能算法^[3, 6-8]的 SVM 算法被相继提出。群体智能优化算法往往具有很强的全局搜索能力, 但是容易出现早熟、陷入局部最优解等问题。另外, 这些算法本身参数设置也是非常复杂的。因此, 选择参数设置简单并且效果好的启发式算法非常重要。模拟退火算法理论上证明能够以概率 1 收敛于全局最优解。本文针对解决的变压器故障问题对模拟退火算法进行了适应性修改, 实验结果表明, 方法是可行并且有效的。

1 变压器故障诊断

油中溶解气体的组合和含量可以在很大程度上反映电力变

压器的运行状况。基于 DGA 数据的变压器故障诊断方法已成为变压器故障诊断的一种有效方法。工业生产中的电力变压器多为充油式变压, 目的是为了绝缘和散热。电力变压器油是由许多不同分子量的碳氢化合物分子组成的混合物, 分子中含碳氢化学基团。在热或电的作用下, 可以使 C-H 和 C-C 键断裂, 同时形成氢气和低分子烃类气体, 如甲烷 (CH₄)、乙烷 (C₂H₆)、乙烯 (C₂H₄)、乙炔 (C₂H₂) [9] 等。电力变压器故障主要分为热性故障和电性故障。其中, 热性故障根据温度分为低温过热, 中温过热, 高温过热 3 类。而电性故障根据放电能量密度分为低能放电, 高能放电以及局部放电。

2 支持向量机

支持向量机是在统计学习理论上发展起来的一种新的机器学习方法, 它建立在结构风险最小化原则基础之上, 在解决有限样本的数据分类时具有很强的学习能力和泛化能力。由于支持向量机在函数表达能力、推广能力和学习效率上都要优于传统的学习机器, 已经被成功地应用于模式识别、回归分析和函数估计等领域, 尤其是在模式识别领域得到了广泛的应用^[10]。

利用核函数实现非线性数据可分是支持向量机最显著的特点之一。核函数本质上是对应于高维空间中的内积的, 从而与生成高维空间的特征映射一一对应。核方法正是借用这种对应关系隐性的使用了非线性特征映射。这一方法不仅使得能够利用高维空间处理非常复杂的数据, 不可分的变成可分的, 同时又回避了高维空间带来的维数灾难。不同的核函数带来不同的

收稿日期: 2014-10-15; 修回日期: 2015-03-31。

作者简介: 谢荣斌 (1970-), 男, 工程硕士, 工程师, 主要从事高电压技术与管理方向的研究。

性能，以及使用不同的参数。因此，核函数是支持向量机解决实际应用中必须面临的一个选择。

在 SVM 训练阶段，并不是训练误差越小越好，甚至会出現误差越小，预测的准确率越低。因此，SVM 引入了松弛变量，即允许离群点 (outlier) 存在误差，而使得整体的分类间隔最大。平衡这种损失以及分类间隔最大就需要引入惩罚因子。

2.1 最大间隔分类器

如果用 x 表示数据点，用 y 表示类别 (y 可以取 1 或者 -1，分别代表两个不同的类)，一个线性分类器的学习目标便是要在 n 维的数据空间中找到一个超平面 (hyper plane)，这个超平面的方程可以表示为：

$$f(x) = w^T x + b \tag{1}$$

对于线性可分的数据而言，存在多个超平面能够分开两类数据。而一个好的分类器的目标就是寻找最适合分开两类数据的平面，即两边的数据的距离这个平面的间隔最大，如图 1 所示。

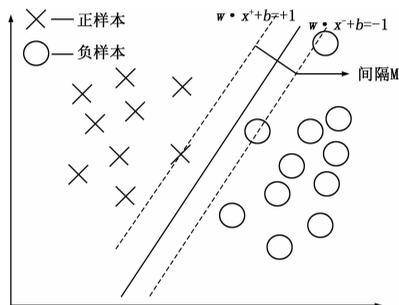


图 1 最大间隔分类

分别在正负样本的平面上取垂直点，记为 x^+ 和 x^- ，则它们两者之间的几何距离就是平面的间隔 M 。

$$\begin{aligned} w \cdot x^+ + b &= +1 \\ w \cdot x^- + b &= -1 \\ \|x^+ - x^-\| &= M \end{aligned} \tag{2}$$

根据公式 (2) 很容易推导出 $M=2/\|w\|$ ，问题转化为带约束条件的最大值问题，等价于求下面的最小化问题。

$$\begin{aligned} \min \frac{1}{2} \|w^2\| \\ y_i (w^T x_i + b) \geq 1, i = 1, \dots, n \end{aligned} \tag{3}$$

上述式子可以通过拉格朗日对偶性变换到对偶变量的优化问题，即通过求解与原问题等价的对偶问题得到原始问题的最优解，这就是线性可分条件下支持向量机的对偶算法，这样做的优点在于：一是对偶问题往往更容易求解；二是可以自然的引入核函数，进而推广到非线性分类问题。

$$L(w, b, \alpha) = \frac{1}{2} \|w^2\| - \sum_{i=1}^n \alpha_i (y_i (w^T x_i + b) - 1) \tag{4}$$

最后利用 SMO 算法求解对偶问题中的拉格朗日乘子。

2.2 SVM 参数

在线性不可分的情况下，支持向量机首先在低维空间中完成计算，然后通过核函数将输入空间映射到高维特征空间，最终在高维特征空间中构造出最优分离超平面，从而把平面上本身不好分的非线性数据分开。常见的核函数有如下几种：

- 1) 线性核函数: $K(x_i, x_j) = x_i^T x_j$
- 2) 多项式核函数: $K(x_i, x_j) = (\gamma x_i^T x_j + r)^d$

- 3) 径向基核函数: $K(x_i, x_j) = \exp(-\gamma x_i - x_j^2)$
 - 4) Sigmoid 核函数: $K(x_i, x_j) = \tanh(\gamma x_i^T x_j + r)$
- 其中: γ, r, d 是核函数的参数。

实际上，径向基核函数 (radial basis function, RBF) 是使用最频繁的。林智仁指出线性核是 RBF 核的特例，而 Sigmoid 核在某些参数下等价于 RBF 核，多项式核的参数多且当 d 很大的时候会出现异常情况。

核函数解决了线性不可分的情况，而分类中还存在着一种过拟合 (overfitting) 的问题。如果要求分割超平面完全的分开所有的数据点，那么当有少数数据点偏离了自身的平面范围，这时候就会出现过拟合的现象。因此，SVM 引入了惩罚因子 C ，目的是为了寻找间隔最大的分类超平面以及保证数据点偏差量最小，如图 2 所示。

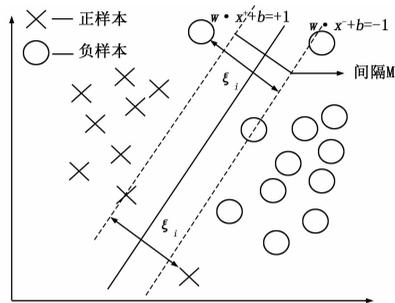


图 2 加入松弛变量

$$\begin{aligned} \min \frac{1}{2} \|w^2\| + C \sum_{i=1}^n \xi_i \\ y_i (w^T x_i + b) \geq 1 - \xi_i, i = 1, \dots, n \\ \xi_i \geq 0, i = 1, \dots, n \end{aligned} \tag{5}$$

其中: C 是固定的常量。同理，利用拉格朗日算子及偶性求解上述问题。

2.3 SVM 多分实现

OVA 是最早用于实现多分的方法。对于具有 $k (k > 2)$ 个类的数据集，将其中一类作为正样本，其余 $k-1$ 类视为负样本，即训练出了 k 个 SVM 模型。在对测试数据集使用的时候， k 个模型中决策函数输出的最大值 x 即为预测的标签。

OVO 的解决方法是对 k 个类进行两两组合，构造 $k(k-1)/2$ 个二分类器，每个分类器只针对两类数据进行训练。在预测的时候通常采用投票法，即如果属于某个类，则将该类的票数加 1，最终得票数最多的类就是预测的结果。

DAG 方法是通过将所有分类器构成 $k(k-1)/2$ 个节点以及 k 个叶子的有向无环图。在分类的时候，首先从根节点开始，根据根节点的分类结果决定下一层选择的节点，直到达到叶子节点为止，该叶子节点所表示的类别即为未知样本的类别。

3 模拟退火算法

模拟退火算法 (simulated annealing algorithm, SAA) 是 20 世纪 80 年代初期发展起来一种求解大规模组合优化问题的随机算法。它是模拟物理系统退火过程来求解全局优化问题。最早由 Metropolis 等于 1953 年提出，随后由 Kirkpatrick 等^[1] 成功地将其思想应用到组合优化领域。模拟退火算法的基本思想是从选定的初始解开始，在借助于控制参数 t 递减时产生的

一系列 Markov 链中, 从当前解产生相邻解, 并利用 Metropolis 接收准则进行判断, 迭代更新当前解, 从而使目标函数最终达到最优。模拟退火算法最大的优点在于能够以一定的概率接收较差解, 以此跳出局部最优的陷阱。假设系统当前状态为 i , 系统能量为 E_i , 对每一个模型参数施加扰动, 产生状态 i 邻域的状态 j , 对应的系统能量为 E_j , 则从状态 i 迁移到状态 j 的概率需要利用如公式 (6) 进行计算。

$$P(S_j) = \begin{cases} \exp\left(-\frac{\Delta E}{T}\right) & (\Delta E < 0) \\ 1 & (\Delta E > 0) \end{cases} \quad (6)$$

借助在 (0-1) 区间的随机数 rand 实现有条件的承认, 如果 P 大于或等于 rand, 则迁移到新状态。从公式可以看出, 温度越高, 接受能量下降的概率就越大; 温度越低, 则接收能量下降的概率就越小。随着温度不断下降, 最终只能向能量增大的状态转移。

模拟退火算法的性能依赖于良好的邻域结构以及退火策略^[12]。

4 自适应 SA 的参数寻优

作为优化问题, 首先需要确定的目标函数; 其次确定 SA 的邻域构造。

4.1 目标函数

SVM 本质上是一个二分类解决方案, 本文主要利用 OVA 方法实现 SVM 的多分类。所有的二分类器共享参数 C 和 γ 。

$$\operatorname{argmax}_{x_i=1, \dots, k} ((w^i)^T \varphi(x) + b^i) \quad (7)$$

公式 (7) 预测的结果作为目标函数的值。

4.2 冷却进度表

模拟退火算法需要确定的参数有初始温度 T_0 , 温度下降率, 以及马尔可夫链链长度 L 。为了避免引入过多的参数以及增加算法的适应性, 初始温度由算法自适应的选择。

$$P = \exp\left(-\frac{\Delta E}{T_0}\right) > 0.9$$

温度的更新策略采用指数下降函数 $T = * T_0$, 其中 (通常设置为 0.9 到 1 之间的数。马尔可夫链长度 L 根据设置的邻域搜索粒度 d 进行动态调整, k 为迭代次数。

$$L = \frac{\log(k)}{d}$$

4.3 邻域解的构造

由于解的维度为 2, 因此本文提出了一种 4 邻域的构造方法, 即每次随机选取上下左右四个直接相邻的解作为其相邻解。 C 和 γ 坐标轴分别代表两个参数的寻优尺度, 即每次参数变化的大小。粗粒度搜索则表示, 每次产生较大的变化。

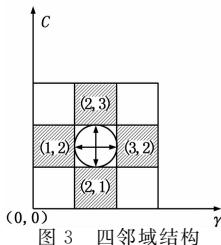


图 3 四邻域结构

4.4 算法流程

如图 4 所示, 算法求解步骤如下。

Step 1: 利用大间距网格搜索估计 SVM 参数;

Step 2: 将 Step1 的参数作为当前解, 并初始化模拟退火的初始温度, 下降率, 迭代结束的标准设为 10 次迭代结果没有发生变化。

Step 3: 将当前解的邻域解作为新解, 并做交叉验证。根据 Metropolis 准则判断是否接收新解, 接收则将新解作为当前解。如果新解的结果优于当前最优值, 则记录当前最好解。重复 Step3 L 次。

Step 4: 下降温度;

Step 5: 判断迭代是否结束; 否, 则重复 Step3, 4; 是, 则将记录的最优解用于预测阶段。

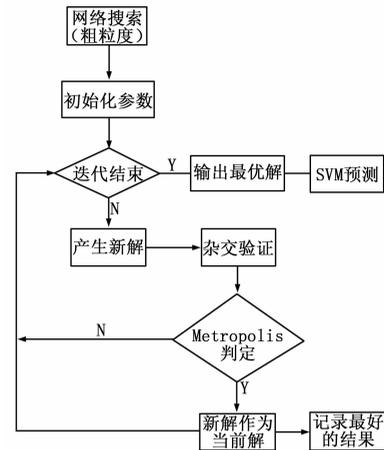


图 4 参数寻优流程图

5 实验结果

根据文献 [9] 提供的数据, 将变压器故障类型标识为六类: 正常, 低能放电, 高能放电, 中低温过热, 高温过热, 局部放电。本文总共搜集了 400 组 DGA 数据, 按各类所占比例选出 320 组作为训练数据, 80 组作为测试数据。本文分别实现了 ASA 方法, 文献 [7] 的 SA-PSO (记为 MGA) 以及文献 [8] 的改进 GA (记作 MGA), 对寻优的结果进行了比较。实验环境为 CPU Intel Q8300、4G 内存、Windows7 操作系统。

表 1 数据集故障类型分布情况

故障类型	正常	低能放电	高能放电	中低温过热	高温过热	局部放电
样本数	52	66	112	43	119	16

本文采用公式 (6)^[3] 对数据进行归一化, 其中 $x_i = [x_{i1} x_{i2} x_{i3} x_{i4} x_{i5}]$ 分别代表 H_2 、 CH_4 、 C_2H_5 、 C_2H_4 、 C_2H_2 气体的体积。对于 H_2 气体, 取其体积分数占 5 种气体总体积分数的百分比数值; 对于烃类气体, 取其体积分数占总烃体积分数的百分比数值。

$$x_{ik}^* = \begin{cases} x_{ik} / \sum_{k=1}^5 x_{ik}, & k = 1 \\ x_{ik} / \sum_{k=2}^5 x_{ik}, & k = 2, 3, 4, 5 \end{cases} \quad (8)$$

网格搜索 (grid search, GS) 代表的 SVM 的自带参数寻优的方法, c 和 r 的范围为 $[-20 \ 20]$ 。SA-PSO 的种群规模为 50, 学习因子为 1.5, 惯性权重 0.4~0.9, 平衡因子 0.06; MGA 的种群规模设为 80, 交叉概率设为 0.9, 变异概率取为 0.1; ASA 的退火率 ($=0.95$, 邻域粒度 $d=0.01$)。SVM 使用

的核函数为高斯核函数。

从表 2 可以看出, ASA 算法和 SA-PSO 算法的准确率比较接近, 明显高于其他两种方法。而 SA-PSO 算法参数多, 运行时间长。

表 2 不同方法的诊断率

方法	ASA	MGA	SA-PSO	GS
诊断正确率/(%)	84.3	82.1	83.9	77.4
参数寻优时间/s	384	430	729	108

表 3 ASA 测试样本诊断结果

故障类型	测试样本数	预测正确数
正常	15	13
低能放电	20	16
高能放电	14	13
中低温过热	10	8
高温过热	15	13
局部放电	6	4

6 结论

模拟退火算法具有坚实的数据理论基础, 能够避免局部最优解的陷阱。将其用于支持向量机的参数寻优, 进行了适应性改进, 避免引入更多的参数设置, 具有更强的适应性, 能够有效地提高故障诊断的正确率。

(上接第 1494 页)

4 试验结果与分析

经过调试运行, 该系统工作稳定, 得到的经纬度信息显示格式如下:

```
<时间><温度><WD: 纬度><JD: 经度>
<11: 14: 27><30.75℃><WD: 38°00.84800 ><
JD: 112°26.54139>。
```

并能将运输车辆的 GPS 信息传输到监控端, 并调用 Google Earth 软件进行定位, 系统运行十分钟后定位精度能保持在 5 m 以内。一旦检测到温度湿度等信息不在正常范围内有发生事故的可能, 就进行声光报警并向目标手机准确发送位置和监控信息, 便于监控中心准确找到有安全隐患的运输车辆。

5 结论

本文提出的移动危化品远程监测系统采用了 MEMS 传感器、LEA-5H 来监测危化品运输的实时状况参数, 弥补了以往定位系统体积大的缺陷, 实现了监测系统的小型化; 使用 ZigBee 实现运输车辆内数据采集模块和移动端主体的通信, 自动报警减少了系统对人力的依赖, 便于事故的预防和处理; 使用 GSM 短信实现运输车辆和监控中心的通信, 点对点的通信使系统有更强的准确性和抗干扰性。本文的研究对危化品的运输监测有重要的意义。

参考文献:

[1] 康 昊, 谭秋林, 秦 丽, 等. 基于 GPS 的车辆安全监测系统 [J].

参考文献:

[1] 熊 浩, 李卫国, 畅广辉, 等. 模糊粗糙集理论在变压器故障诊断中的应用 [J]. 中国电机工程学报, 2008 (7): 141-147.

[2] 刘守明, 胡志坤, 王美玲. 基于知识库的电力变压器故障诊断专家系统 [J]. 计算机测量与控制, 2011 (7): 1569-1571.

[3] 郑蕊蕊, 赵继印, 赵婷婷, 等. 基于遗传支持向量机和灰色人工免疫算法的电力变压器故障诊断 [J]. 中国电机工程学报, 2011 (7): 56-63.

[4] 徒 伟, 张广明. 基于支持向量机的多类分类在变压器故障诊断中的应用 [J]. 计算机应用, 2010 (A1): 97-98.

[5] 董 明, 孟源源, 徐长响, 等. 基于支持向量机及油中溶解气体分析的大型电力变压器故障诊断模型研究 [J]. 中国电机工程学报, 2003 (7): 88-92.

[6] 肖燕彩, 陈秀海, 朱衡君. 遗传支持向量机在电力变压器故障诊断中的应用 [J]. 上海交通大学学报, 2007 (11): 1878-1881.

[7] 郑 凯, 袁 丹, 刘剑清, 等. 基于 SA-PSO 优化自适应 PNN 网络的变压器故障诊断研究 [J]. 计算机测量与控制, 2014 (4): 1015-1017.

[8] 王建玺, 王刘涛. 基于改进 GA 的 SVM 电力变压器过热诊断方法研究 [J]. 计算机测量与控制, 2014 (2): 342-344.

[9] 尹金良. 基于相关向量机的油浸式电力变压器故障诊断方法研究 [D]. 北京: 华北电力大学, 2013.

[10] 丁世飞, 齐丙娟, 谭红艳. 支持向量机理论与算法研究综述 [J]. 电子科技大学学报, 2011 (1): 2-10.

[11] S. Kirkpatrick. Optimization by simulated annealing; Quantitative studies [J]. Journal of statistical physics 1984, 34, 975-986.

[12] 傅文渊, 凌朝东. 布朗运动模拟退火算法 [J]. 计算机学报, 2014 (6): 1301-1308.

计算机测量与控制, 2013, 21 (9): 2406-2408.

[2] 朱向庆, 陈志雄. 采用 GPS 和 GPRS 的移动目标监控导航系统设计 [J]. 计算机测量与控制, 2011, 19 (5): 1153-1157.

[3] 童长飞. C805 系列单片机开发与 C 语言编程 [M]. 北京: 北京航空航天大学出版社, 2005.

[4] 李 耀, 昂志敏, 李敏杰, 等. 基于 3G 车载移动终端的 GPS 定位系统设计 [J]. 微型机与应用, 2012, 31 (23): 51-54.

[5] 程 铭. 基于 GPS+GSM 双技术定位的车辆跟踪系统的研究 [D]. 天津: 天津大学, 2010.

[6] 韩英梅, 赵建平, 白慧静. 基于 DHT11 的无线温湿度传感器网络节点的设计 [J]. 井冈山大学学报 (自然科学版), 2011 (1): 67-70.

[7] 袁 西, 陈 栋, 田 湘, 等. 三轴数字加速度计 ADXL345 及其在捷联惯导中的应用 [J]. 国外电子元件, 2010 (3): 138-140.

[8] 高 锋, 谭春毅. 新型双向加速度传感器 MMA6260 的原理与应用 [J]. 传感技术学报, 2005, 18 (1): 150-152.

[9] 谭秋林. MEMS 红外瓦斯传感检测系统的研究 [D]. 太原: 中北大学, 2006.

[10] 王泽民, 徐绍拴, 等. GPS 测量原理及应用 [M]. 武汉: 武汉大学出版社, 2008.

[11] 漆明生, 徐家清. 基于 KML 的地理数据表示 [J]. 四川林勘设计, 2008 (1): 77-80.

[12] 杨 雪, 黄 坚, 翁敬农. KML 及其在“数字校园”中的应用实现 [A]. 中国地理信息系统协会第四次会员代表大会暨第十一届年会论文集 [C]. 2007: 690-697.

[13] 刘 冰, 石奉华. Google Earth 在旅游、导航中的应用探讨 [J]. 山东科技大学学报 (自然科学版), 2006, 25 (4): 25-28.