

混合训练的 DHMM 及其在发射机状态检测中的应用

许丽佳^{①②} 龙兵^① 王厚军^①

^①(电子科技大学自动化工程学院 成都 610054)

^②(四川农业大学信息与工程技术学院 雅安 625014)

摘要: 隐马尔可夫模型(HMM)是一种双随机过程,其训练方法 B-W 算法是一种基于爬山算法,容易陷入局部最优且对初始参数值依赖性大。为了提高模型的有效性,该文提出了将改进的模拟退火(SA)算法和 B-W 算法相结合的混合训练方法,解决了受模型参数初值影响的问题并能实现全局搜索。将其应用于发射机功率状态检测中,实验结果证明该方法准确性高,收敛速度快和稳定性好,是一种很有实用价值的新方法。

关键词: 隐马尔可夫模型; 模拟退火算法; KL 距离

中图分类号: TN911.7, TN06

文献标识码: A

文章编号: 1009-5896(2008)07-1661-05

Hybrid Training DHMM and Its Application to Check Transmitter Power

Xu Li-jia^{①②} Long Bing^① Wang Hou-jun^①

^①(Institute of Automation Engineering, UEST of China, Chengdu 610054, China)

^②(Institute of Information & Engineering Technology, Sichuan Agriculture University, Yaan 625014, China)

Abstract: HMM model is a double random processing which is trained with B-W algorithm, this algorithm based on hill-climbing is easy to lead to locally optimal solutions and its validity is greatly depend on model initial parameters. In order to improve the validity of model, this paper proposes a hybrid training method which combine the B-W algorithm with improved SA algorithm. With this hybrid method the validity of model is not influenced by the model initial parameters and the global optimal solution can be easily gained. Applying this hybrid method to check transmitter power, the experimental results show that proposed method is practical method and own qualities such as high veracity, rapid converged and good stability.

Key words: Hidden Markov Model (HMM); SA algorithm; Kullback distance

1 引言

隐马尔可夫模型 HMM(Hidden Markov Model)是一种广泛用于信号处理和模式识别中的双随机过程,具有完善的数学理论基础和良好的模式分类能力,已经成功地应用于语音识别领域^[1],也开始逐渐应用于故障诊断领域^[2]。常规的 HMM 模型根据观测信号的性质分为两类:连续隐马尔可夫模型 CHMM 和离散隐马尔可夫模型 DHMM。该模型主要解决 3 类问题:(1) 评估问题:计算由给定模型 λ 产生观测序列 O 的概率 $P(O|\lambda)$;(2) 解码问题:对给定模型 λ 和观测序列 O ,求可能性最大的状态序列 S ;(3) 学习问题:对给定观测序列 O 在最大似然度下学习得到模型 $\lambda = (A, B, \pi)$: $\max P(O|\lambda)$ 。在这 3 类问题中,其中最重要的就是 HMM 模型的参数训练问题,常规的 Baum-Welch(B-W)算法是一种基于爬山的算法,它的基本思想就是在搜索过程中,每次迭代都会朝着似然概率增大的方向,因此具有收敛速度快的优

点,但同时有一个致命弱点:容易陷入局部最优值。该算法对模型初始参数依赖很大,常常训练出的参数却只是初始参数的邻近值。同样的样本在随机给初始参数训练的情况下得到的似然概率变化也很大,所以这使得 HMM 模型识别能力大大下降,对训练算法的改进很有必要。现在有提出一种 K 均值分段与聚类算法^[3]来对初始参数作初步估计,也有提出梯度下降训练^[4]或遗传算法来训练^[5, 6],计算量复杂。

模拟退火算法(SA)^[7]是一种基于全局搜索的类似于固体的退火过程的算法,最早是由 Metropolis 在 1953 年提出,由 Kirkpatrick 等人在 1983 才将其成功用于求解组合优化问题。由于模拟退火算法是全局搜索最优解,并且理论上它要求初始温度必须足够高,每个温度下状态交换必须足够充分,温度下降必须足够慢的前提,这就决定了该算法收敛速度慢,计算量大的特点。

本文受此启发,提出将局部寻优且收敛速度快的 B-W 算法与全局寻优收敛速度却很慢的 SA 算法相结合的混合训练算法,使得训练出的 DHMM 模型兼备识别能力强,准确性高且收敛速度快的特点。最后将其应用于雷达故障状态检

2006-11-27 收到, 2007-05-21 改回

国家部级项目和电子科技大学博士台建设项目(05BS00701)资助课题

测, 实验结果证明其有效性。

2 HMM 模型及其训练算法

隐马尔可夫模型 HMM 是一种双重马尔可夫随机过程, 它包括具有状态转移概率阵 \mathbf{A} 的 Markov 链和输出观测值的随机过程, 其状态是不确定或不可见的(即隐藏的), 只有通过观测序列的随机过程才能表现出来(即通过观测概率矩阵 \mathbf{B} 相联系)。它常用五元组来描述: $(\Omega x, \Omega o, \mathbf{A}, \mathbf{B}, \pi)$ 。 $\Omega x = \{q_1, \dots, q_N\}$ 为状态的有限集合; $\Omega o = \{v_1, \dots, v_M\}$ 为观测值的有限集合; $\mathbf{A} = \{a_{ij}\}, a_{ij} = P(x_{t+1} = q_j | x_t = q_i)$ 为状态一步转移概率阵; $\mathbf{B} = \{b_{ik}\}, b_{ik} = P(o_t = v_k | x_t = q_i)$ 为输出概率阵; $\pi = \{\pi_i\}, \pi_i = P(x_1 = q_i)$ 为初始状态分布概率。常将 HMM 模型简记为 $\lambda = \{\pi, \mathbf{A}, \mathbf{B}\}$ 。采用前向-后向算法来计算由模型 λ 产生观测序列 O 的概率, 用 Viterbi 算法来计算由模型 λ 和给定观测序列 O 确定最佳状态序列及其概率, 用 Baum-Welch 算法通过观测序列在最大似然估计概率意义下训练得到的模型 λ 。由此可以看出, 模型 λ 的值适合与否会严重影响前两个概率的计算, 即影响 HMM 模型识别能力。HMM 模型问题主要集中在模型 λ 的训练算法上。

Baum-Welch 算法的思想是先任意给定一个初始模型 λ_0 , 基于 λ_0 以及观测序列 O , 通过前向-后向算法计算得到 $P(O | \lambda_0)$, 然后通过 EM 算法进行模型的训练得到新的模型参数 λ , 若 $P(O | \lambda) - P(O | \lambda_0) < \text{阈值}$ (收敛条件), 说明训练已经得到预期效果, 算法结束, 否则就令 $\lambda = \lambda_0$, 进行反复训练一直到收敛为止。本文为了增加 HMM 模型的稳定性, 采用多组观测序列组进行训练, 得到的参数估计公式如下:

$$\left. \begin{aligned} \bar{\pi}_i &= \frac{\sum_{l=1}^L \alpha_{1,l}(i) \beta_{1,l}(i) / P(O_l | \lambda)}{\sum_{l=1}^L \sum_{t=1}^{T_l-1} \alpha_{t,l}(i) a_{ij} b_j(O_{t+1,l}) \beta_{t+1,l}(j) / P(O_l | \lambda)} \\ \bar{a}_{ij} &= \frac{\sum_{l=1}^L \sum_{t=1}^{T_l-1} \alpha_{t,l}(i) \beta_{t+1,l}(j) / P(O_l | \lambda)}{\sum_{l=1}^L \sum_{t=1}^{T_l-1} \alpha_{t,l}(i) \beta_{t+1,l}(j) / P(O_l | \lambda)} \\ \bar{b}_{jk} &= \frac{\sum_{l=1}^L \sum_{t=1 \cap O_t=V_k}^{T_l-1} \alpha_{t,l}(j) \beta_{t+1,l}(j) / P(O_l | \lambda)}{\sum_{l=1}^L \sum_{t=1}^{T_l-1} \alpha_{t,l}(j) \beta_{t+1,l}(j) / P(O_l | \lambda)} \end{aligned} \right\} \quad (1)$$

其中 α 为前向变量; β 为后向变量; L 为观测序列的组数; T_l 为每组观测序列长度。

由于在 Baum-Welch 算法中, 每次迭代过程都是爬山过程, 使得最终收敛到局部最小值。所以本文尝试引入改进的模拟退火算法来训练 HMM 模型。

3 改进的 SA 算法

3.1 SA 算法原理^[7]

模拟退火算法(SA)是根据复杂组合优化问题与固体的退火过程之间的相似性提出来的。固体的退火过程是一种物

理现象。当对一个固体进行退火(即降温)时, 随着温度的下降, 粒子的热运动逐渐减弱并停留在不同的状态, 其排列也从无序向有序发展, 直至温度很低时, 粒子重新以一定的结构排列, 也就对应着不同的能量水平。若退火过程缓慢进行, 使得在每个温度下都得到一种平衡态, 当温度趋于 0 时, 系统的能量将趋于最小值。若以粒子的排列或相应的能量来表达固体所处状态, 给出了从状态 i 转换到状态 j 的准则:

如果 $E(j) \leq E(i)$, 则状态转换被接受; 如果 $E(j) > E(i)$, 则状态转移被接受的概率为

$$P = e^{-\frac{E(i)-E(j)}{kT}}$$

其中 $E(i)$, $E(j)$ 表示状态 i , j 下的能量, k 是波尔兹常数, T 是温度。

3.2 SA 算法的改进步骤

SA 算法按照 Metropolis 准则随机接受一些劣解, 即指标函数大的解。当温度较高时, 接受劣解概率较大, 随着温度下降, 接受劣解概率逐渐减少, 直到温度为 0 时接受劣解概率也趋于 0。正是因为这样, 将有利于算法从局部最优解中跳出来从而搜索到全局最优解。所以要求在退火过程中初始温度足够高(高温下接受劣解概率大, 易跳出局部最优值), 每个温度下状态交换足够充分, 温度下降也应当足够缓慢。全局计算量又非常大, 这就决定了 SA 算法的缓慢性。本文从以下几方面改进以加快算法快速性和搜索到全局最优解的几率:

(1) 温度下降表: 原 SA 算法的温度下降采用公式为 $T_t = T_0 / \log(1+t)$, 而本文采取 $T_t = \mu T_{t-1}$, μ 为略小于 1 的常数。

(2) 每个温度下的迭代次数: 由于要求每个温度下状态交换足够充分, 传统 SA 算法规定在每个温度下的迭代数为比较大的固定值, 导致计算量很大。在高温时比较小的迭代数就可得到平衡态, 而随着温度下降被拒绝的状态数增加, 因此在低温下应当迭代数增加, 避免过早陷入局部最优值。本文设计在每个温度下的迭代次数为可变步长, 与温度近似成反比关系, 这样意味随着温度下降应适当增加迭代次数。

(3) 设置全局最优变量: 设置记忆变量 λ' 和目标函数值 $E(\lambda')$ 用于储存当前的最优解和最优目标函数值。初始时令 $\lambda' = \lambda_0$ 和 $E(\lambda') = E(\lambda_0)$, 以后每搜索到一个新解时将目标函数值 $E(\lambda_k)$ 与 $E(\lambda')$ 相比较, 若 $E(\lambda_k)$ 优于 $E(\lambda')$ 则用 λ_k , $E(\lambda_k)$ 分别取代 λ' 和 $E(\lambda')$, 直至算法结束。这时的 λ' , $E(\lambda')$ 可看作算法的近似全局最优值, 就避免算法陷入局部最优值。

(4) 设置无效退火次数: 典型 SA 算法是从很高的初始温度缓慢下降到允许的最低温度为此。为了加快算法的快速性能, 本文算法设置一个无效退火次数 m , 若在相邻的 n 个温度中得到的目标函数值无任何变化, 就说明算法已经收敛, 即便是收敛于局部最优值, 由于在低温下跳出的可能性也很

小,因此可以终止算法。

4 B-W 和 SA 混合训练算法设计

从上面可知 SA 算法理论上能搜索到全局最优值,可搜索非常缓慢,尤其是局部搜索能力差,尽管有了几种改进措施但不能从根本上解决问题。考虑到 B-W 算法局部搜索能力强,收敛速度快的特点,在 SA 算法的每次迭代中,产生的新解先用 B-W 算法训练后再作为新解进入迭代运算。

4.1 能量函数的选取

由于 HMM 模型的训练是希望训练数据对模型的似然概率越高越好,所以选定目标能量函数为

$$E(\lambda) = -1 \times \sum_{k=1}^R \ln [P(O^{(k)} | \lambda)] \quad (2)$$

也就意味着当 SA 算法搜索到全局 $E(\lambda)$ 最小时,训练数据对模型的似然概率值也就为最大值。

4.2 邻近参数的产生

邻近值的产生是随机的,但其产生方法在一定程度上会影响 SA 算法的收敛效率。文中设计产生邻近值方法如下:

$m'_i = m_i + \eta u$ 。其中 m_i 为当前模型中的第 i 个参数, u 为 $[0, 1]$ 上均匀分布的随机数, η 为扰动幅度。若 η 选取太小,则不易跳出局部最优值,若太大又影响收敛速度。对 A, B 矩阵,每个参数产生扰动后对其进行归一化处理。

4.3 混合算法的具体步骤

混合训练算法具体步骤如下:

(1)初始化模型参数;

(2)由前-后向算法计算出概率似然度值并取反得初始能量值 E -old,并赋给全局最优目标函数值,保存 $E(\lambda')$ $E(\lambda')$ 和模型参数值 λ' ,并设置温度退火次数 $kk=1$;

(3)若温度已经下降到最低温度或无效退火次数已经达到允许的最大次数或总迭代次数 kk 大于允许的最大迭代次数则转步骤(13),否则 $kk=kk+1$,运行下一步;

(4)初始化为 $j=1$,设置每个温度下迭代次数 $L(T_j)=u/T_j$, u 为常数。若 $j < L(T)$,继续下一步。否则就转到步骤(11);

(5)对模型中的所有参数 $\lambda(A, B)$ 产生邻近值,然后进行归一化处理得到参数;

(6)用 B-W 算法对归一化后的参数用多组观察数据进行训练,得训练后的模型新参数 λ -new 和似然概率值,对该似然概率值取反就得到新的能量值 E -new;

(7)比较新能量值与全局最优目标函数值 $E(\lambda')$:若 E -new $<$ $E(\lambda')$,就令 $E(\lambda')=E$ -new, $\lambda' = \lambda$ -new;

(8)比较 E -new 和 E -old:若 E -new $<$ E -old,就接受新状态(E -old= E -new, λ -old= λ -new)否则执行下一步;

(9)计算概率 $P = e^{-\frac{E-new-E-old}{kT}}$,随机产生一个小于 1 的正数 num,若 $P >$ num,就接受新状态,否则就拒绝新状态。 $j=j+1$,转到步骤(5)。

(10)将 E -train(kk)= $E(\lambda')$;

(11)比较 E -train(kk)和 E -train($kk-1$):

若 E -train(kk) $<$ E -train($kk-1$),则无效退火次数 $m=0$,否则 $m=m+1$,继续下一步;

(12)进行温度下降: $T_j = \mu T_{j-1}$,并转到步骤(3);

(13)算法结束。

5 混合训练的DHMM在发射机状态检测中的应用

由于磁控管电流的变化很大程度上反映发射机功率大小,若电流偏低则意味着雷达灵敏度降低或发射机功率降低,故选择最能反映发射机功率变化的磁控管电流来实现对发射机功率状态的监测。

5.1 混合训练算法精度和快速性分析

以 L 波段雷达为例,正常情况下磁控管电流约在 2.4mA 上下波动,以采样速率 1kHz 进行电流信号采集,一共采集了 7500 点,分为 200 帧(50 点/帧),每帧经过 SSNF 去噪(采用非正交二进小波将信号分解为 6 层后,通过相关去噪算法获得新的小波系数再进行重构得到去噪信号)后进行特征提取,由 'db1' 小波包分 3 层得到 8 频段归一化能量分布图。由于能量主要集中在第一频段,所以就以该频段能量值作为反映发射机功率状态检测的特征值。所以共得 150 个特征能量值,分为 15 组(即每组观测数据长度为 10)观测数据来完成对 DHMM 模型的训练,由于采用 DHMM 模型故对训练数据还作量化处理。实验采用 5 个状态数,20 个观测值的无跳跃的左右型 DHMM 模型。初始状态转移概率取值为: $\pi = [1, 0, 0, 0, 0]$; $A = [0.5, 0.5, 0, 0, 0; 0.5, 0.5, 0, 0, 0; 0, 0, 0.5, 0.5, 0; 0, 0, 0.5, 0.5; 0, 0, 0, 1]$,观测阵 B 采取随机取值。分别用传统 B-W 算法(迭代到概率似然度值变化小于 $1e-5$ 为止)和文中混合算法得到的 15 组数据所对应的概率似然度值。由于 p, A 参数初值选取影响不大,所以在实验中随机选取 10 次 B 阵的不同初值。在混合训练算法中设置 B-W 算法迭代 10 次,连续无效退火最大允许次数为 15 次,设置初始温度为 1000,温度下降系数为 0.95,允许最大迭代次数为 60, $\eta = 1$ 。表 1 列出 10 次不同 B 阵初值所训练出的模型对数似然度值(各表数据全为似然度取对数)。

从表 1 数据看出 B 阵初值的随机选择对 B-W 算法影响很大,相同样本训练出的 DHMM 模型所对应的似然概率值波动较大,这说明原算法依赖于 B 阵初值的选取,对比可知混合训练算法训练出的 DHMM 模型稳定性要高得多,并且对应的 10 次平均似然概率值 -265.8808 要远高于用原算法训练得到的平均似然概率值 -281.3658,即准确度明显增加。

图 1 中虚线为混合算法训练曲线,实线为原算法训练图。由于混合算法加入了对传统 SA 算法的快速性改进措施,从表 1 中迭代次数可以看出该算法迭代次数也不大,从程序运行时间上看,由于在混合训练算法中每次迭代包含了 B-W 算法 10 次,所以混合算法的一次迭代要相当于原算法 10 次迭代(图 1 中混合算法的迭代次数就是用表 1 中迭代次数乘

10 得到), 即便如此改进算法程序运行时间比原算法运行时间平均多 10^5 左右, 而这完全可以接受, 说明混合算法程序收敛性也好。

表1 两种方法训练结果(似然度取对数)比较

次数	混合算法	迭代数	B-W 算法	迭代数
1	-265.6488	14	-289.2481	42
2	-268.5678	8	-273.8709	23
3	-263.1912	18	-282.2159	39
4	-268.6593	3	-287.9314	33
5	-268.1948	26	-284.1378	29
6	-263.7652	43	-279.2793	22
7	-268.2900	7	-274.6260	20
8	-264.5358	31	-271.5811	27
9	-264.0154	19	-282.0105	28
10	-263.9393	24	-288.9566	25
平均值	-265.8808	19.3	-281.3858	28.8

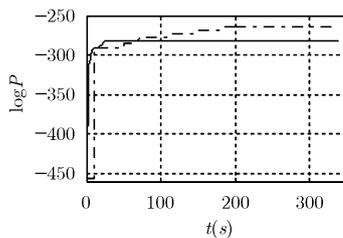


图1 B-W法与混合法训练比较图

为了研究快速性和准确性间的关系, 在混合算法中设置内迭代 15 次, 连续无效退火最大允许次数为 20 次, 其余参数不变。随机选取 B 阵初值 15 次进行训练, 得到结果模型似然度值见表 2。由表 2 可计算 15 次实验平均似然概率值为 -262.1579, 平均迭代次数为 42.3。与表 1 比较可知 DHMM 模型精度提高的同时程序运行时间也增加, 尤其是在温度较低时, 由程序设计中的可变步长得知在低温上达到平衡的时间就需要更多。

从图 2 看出(虚线为混合法训练曲线, 实线为 SA 法训练曲线), 若只用 SA 算法训练, 由于局部搜索能力差, 计算量很大, 程序运行非常缓慢, 当迭代 1000 次后, 混合算法早

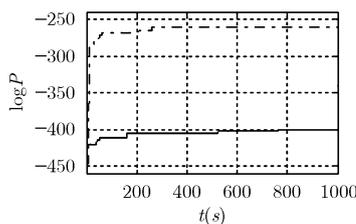


图2 纯 SA 法与混合法训练比较图

就收敛在高似然度值-260.28, 而 SA 算法上升很慢, 且似然度值为-400.26。在理论上只要初始参数取得恰当, SA 算法最终会收敛到全局最优值。

由图 1, 图 2, 表 1 数据可知, 混合算法集中了两单独算法优点, 精度高于任一单独算法且收敛速度快。

5.2 DHMM 在发射机状态检测中的应用

定义 KL 距离: 设 p 是问题域 U 上的一个概率密度函数, 另一个概率密度函数为 q , q 是 p 的近似, 则 q 和 p 之间的距离定义为

$$d_{KL}(p || q) = \sum_{x \in U} p(x) \log \frac{p(x)}{q(x)} \quad (3)$$

KL 距离的大小表示了 p 和 q 之间的接近程度, 若 KL 距离越小, 表示 p 越近似于 q , 当且仅当 $p=q$ 时, KL 距离为 0, 由于 p, q 本身太小没有实际物理意义, 所以本文在计算 KL 距离时直接采用表中的对数似然值代替。

状态检测的思想如下: 获取发射机功率正常时所对应的磁控管电流的特征向量作为训练数据, 用混合算法来训练 DHMM 模型并求得 $P_{\text{正常}}(O | \lambda)$, 然后用发射机功率未知状态所对应的磁控管电流特征量, 通过训练好的 DHMM 模型计算出对应的 $P_{\text{未知}}(O' | \lambda)$, 由于程序中数据产生有一定随机性, 故对 B 阵随机取初值 n 次, 就可得 n 对 $[P_{\text{正常}}(O | \lambda), P_{\text{未知}}(O' | \lambda)]$, 从而计算出 KL 距离值, 根据其值大小就可判别出未知状态偏离正常状态的程度, 若 KL 值很小说明此时设备状态正常, 随着 KL 值的增大发射机出现故障概率也增加, 若大于阈值故障已经出现, 从而实现了发射机故障预测作用。

分别采集发射机功率正常时磁控管电流 7500 个点和未知状态时的磁控管电流 2500 个点, 经过去噪与特征提取后获得正常状态时 15 组训练数据和 5 组测试数据, 用正常态的 15 组数据分别通过混合算法和原 B-W 算法训练作状态检测的两个 DHMM 模型。再由训练好的两个 DHMM 模型分别计算测试数据的似然概率值, 实验结果见表 3(注: 为了与 5 组测试数据对应比较, 其中训练似然数据是 15 组训练数据得出的 5 组平均似然值之和)。

由实际情况设定出现故障时所对应的 KL 距离阈值, 根据实际 KL 值接近阈值的程度来估计发射机故障可能性大小并估算故障将出现的时间, 从实际的不易判断并淹没在噪声中的蕴涵设备故障状态变化趋势的微弱信息转变为明显就能判断的 KL 距离变化值。从表 3 中数据可以看出混合法的稳定性和模型的精度(或有效性)都比原算法好的多。第 3 列和第 5 列的数据是测试样本经两种算法训练出的 DHMM 模型(在相同初始模型情况下)计算出的似然概率值, 可知直接用原算法对发射机状态进行估计是很不稳定的, 得到的似然概率波动很大, 容易因模型初值的随机性而出现对状态估计的完全相反结果, 如第 3 组数据(对应 KL 值为 97.482)易判断发射机已出现故障而第 9 组数据(对应 KL 值为 1.717)却易

表2 混合算法运行结果(似然度取对数)

次数	1	2	3	4	5
似然	-262.6913	-263.0787	-260.4551	-260.8389	-260.2445
迭代	27	38	48	47	33
次数	6	7	8	9	10
似然	-264.6657	-260.2975	-261.8054	-262.1290	-264.4247
迭代	36	60	37	60	50
次数	11	12	13	14	15
似然	-260.2496	-261.2584	-266.5745	-261.5265	-262.1288
迭代	60	37	41	60	60

表3 B-W法与混合法测试结果比较

次数	混合法训练似然值	混合法测试似然值	B-W法训练似然值	B-W法测试似然值
1	-91.9672	-325.9638	-94.8996	-743.1896
2	-92.0833	-321.9171	-99.7865	-106.7526
3	-91.9625	-355.5421	-94.4204	-1017.400
4	-91.9686	-300.7983	-92.0680	-718.4390
5	-92.1063	-337.5984	-96.8219	-270.7931
6	-92.0873	-325.1777	-93.3962	-391.8351
7	-92.1061	-329.9304	-92.1622	-615.9153
8	-92.1103	-343.0965	-95.4507	-245.6268
9	-91.9598	-340.8776	-96.7857	-100.8227
10	-91.5650	-324.7329	-97.0267	-349.1944
11	-92.1264	-313.9090	-92.6485	-580.6307
12	-91.9573	-351.0136	-93.3720	-773.1803
13	-92.0952	-319.8059	-96.0270	-111.6986
14	-92.0918	-326.5372	-97.1225	-890.2103
15	-92.1013	-321.6273	-95.4233	-304.1661
均似然值	-92.0192	-329.2352	-95.1607	-481.3236
KL距离(15次)		763.6358	847.0506	
KL距离(1次)		50.9090	56.4700	

得出发射机还完全正常的结论。而混合算法训练的DHMM模型计算出的测试数据似然值就稳定的多,不会出现误判结果。

6 结束语

本文分析DHMM模型原训练算法的不足,提出了将基于全局搜索的模拟退火算法与基于爬山的改进B-W算法相

结合,充分利用各自的优点,弥补它们的缺陷,实验结果充分验证了该算法具有准确性高且收敛性好的优点。并把该方法成功应用于发射机功率状态检测过程中,实验结果证明混合训练法的有效性和稳定性,为发射机故障预测提供了新方法和新思路。

参考文献

- [1] Rabiner L R. A tutorial on hidden Markov models and selected applications in speech recognition [J]. *Proc. IEEE*, 1989, 77(2): 257-286.
- [2] Jie Ying, and Kirubarajan T. A hidden Markov-based algorithm for fault diagnosis with partial and imperfect tests [J]. *IEEE Trans. on System Man and Cybernetics*, 2000, 30(4): 463-473.
- [3] Gales M J F. Cluster adaptive training of hidden Markov model [J]. *IEEE Trans. on Speech and Audio Processing*, 2000, 8(4): 417-432.
- [4] 茅晓泉, 胡光锐. 一种基于梯度的HMM参数重估方法[J]. *上海交通大学学报*, 2002, 36(5): 683-685.
- [5] Chau C W, Wang S K, and Diu C K. Optimization of HMM by a genetic algorithm [C]. *Proc. Munich, Germany*, 1997: 1727-1730.
- [6] 何志国, 何钦铭, 陈奇. 语音识别中进化计算与MDI的HMM训练方法 [J]. *系统工程理论与实践*, 2006, 26(7): 54-58.
- [7] 孙延奎, 朱小燕. *人工智能[M]*. 第一版. 北京: 清华大学出版社, 2004: 220-246.

许丽佳: 女, 1973年生, 讲师, 博士生, 研究方向为信号处理、故障预测与诊断。

龙兵: 男, 1974年生, 讲师, 博士, 研究方向为故障预测与诊断。

王厚军: 男, 1961年生, 教授, 博士生导师, 研究方向为故障预测与诊断。