

一种计算话务量的进化神经网络方法

夏太武¹,刘金祥¹,彭京华²

XIA Tai-wu¹,LIU Jin-xiang¹,PENG Jing-hua²

1.邵阳学院 信息与电气工程系,湖南 邵阳 422000

2.湘华科技信息有限公司,长沙 410001

1.Department of Information and Electrical Engineering,Shaoyang University,Shaoyang,Hunan 422000,China

2.Xianghua Scientific and Technological Information Ltd,Changsha 410001,China

E-mail:xtw2009@126.com

XIA Tai-wu,LIU Jin-xiang,PENG Jing-hua.Method of evolutionary neural computation on traffic of calls.Computer Engineering and Applications,2007,43(30):130-132.

Abstract: The traffic of calls is a very important parameter in degree measurement for telephone equipment frequency used by users.The distribution of the traffic of calls presents the characteristics of tridimensionality,more services and none-Possion flow,in this case,the Erlang B formula can't be applied to compute wireless block ratio directly.Therefore,according to the computational intelligence,a new method of evolutionary neural computation based on Particle Swarm Optimization(PSO) is presented to analyze and forecast the traffic of calls,the main process includes acquiring data and normalizing,such as traffic of calls,optimizing the structure of neural network,training network by PSO,calculating traffic of calls.By researching on service of personal handy phone call about a networks telecommunication company in Hebei province,the calculating model is set up with the sample information, which includes wireless block ratio,call completing ratio,call dropping ratio.The accuracy of the traffic of calls calculated by the model is very high,which shows the effect of the method is remarkable.

Key words: traffic of calls;evolutionary neural computation;Particle Swarm Optimization(PSO);network structure optimization

摘 要:话务量是度量用户使用电话设备频繁程度的一个重要参量,由于目前话务分布呈现出显著的立体性、多业务性和非泊松流等特点,不能直接应用欧兰 B 公式进行计算。为此,从计算智能出发提出一种基于 PSO 算法的进化神经计算方法,主要包括话务量及其相关参量的获取、神经网络结构的优化、基于 PSO 算法的网络训练,以及话务量计算等步骤。通过对河北省某市小灵通业务的详细研究,利用近半年来的话务量与无线阻塞率、来话接通率和掉话率等参量构成的样本信息进行建模,所计算的话务量精度高,表明其方法切实可行且效果显著。

关键词:话务量;进化神经计算;PSO;网络结构优化

文章编号:1002-8331(2007)30-0130-03 **文献标识码:**A **中图分类号:**TP18

1 问题的提出

话务量是度量用户使用电话设备频繁程度的一个重要参量,也是用来对无线通信系统进行评价、预测和决策的一个重要依据。随着数据业务迅猛发展,话务分布特点呈现出显著的立体性和多业务性,在实际通信网络中,话务量是与阻塞率、来话接通率和掉话率等多种参量相关的,因此话务量实际是一个动态的过程,流入信道的话务往往是具有峰值特性的话务流,即非泊松流,不能直接应用欧兰 B 公式^[1]。

目前,计算或预测话务量常采用时间序列预测方法,将话务量的历史数据看成时间序列,利用加权平均数对时间序列进行平滑修匀,再利用数理统计方法进行回归处理,拟合数学模型,根据数学模型预测未来的话务量。但这种方法要求历史数据必须完整,而且不适用于有季节变动规律的月份话务量预

测^[2,3]。为此,本文从计算智能技术出发,提出一种计算话务量的方法,具体工作如下:

(1)提出一种基于 PSO 算法的进化神经计算方法,利用阻塞率、来话接通率和掉话率等多种相关参量来计算话务量。在网络学习训练中,采用再现群智能的 PSO 算法具有简捷快速、全局寻优等优点,克服了传统 BP 及其改进算法的收敛速度慢、精度低、易于陷入局部极小点等诸多缺陷。

(2)采用黄金分割优选法在网络隐节点数“频繁”出现的区间中寻优,再在由黄金分割方法扩展的区间中寻优,最后综合考虑网络的逼近能力和泛化能力,选取理想的隐含层节点数,以达到搜索效率高、节省成本、优化网络结构的目的。

(3)针对河北省某市网通公司提供的阻塞率、来话接通率、掉话率和话务量的小灵通业务数据作为样本信息进行实际处

基金项目:国家自然科学基金(the National Natural Science Foundation of China under Grant No.60377020)。

作者简介:夏太武(1969-),男,讲师,主要从事人工智能的研究;刘金祥(1965-),男,高级讲师,主要从事计算机软件的教学和科研工作;彭京华(1975-),女,湖南大学在职硕士生,主要从事计算智能与软件研究工作。

理, 其话务量的计算精度高。

2 计算方法与系统设计

2.1 基于 PSO 算法的神经网络学习算法

人工神经网络是一种优越于传统统计学的智能信息处理技术, 其网络的逼近能力与学习算法紧密相关。在前向网络的学习中, 常用基于梯度下降法的 BP 算法, 即让参数沿着误差梯度相反的方向移动, 使误差函数减小, 直到取得极小值。由于 BP 算法存在学习速率慢, 常收敛于局部极小值或在误差函数的平稳段停滞不前等缺陷, 国内外作者提出了不少改进算法, 其核心就是调整学习速率、增加惯性量, 但改进的 BP 算法始终未能克服收敛速度慢和局部极小的问题^[4]。为此本文采用再现群智能的进化计算方法——PSO 算法来进行网络训练。

粒子群优化算法(PSO)源于对鸟群捕食的行为研究, 由 Eberhart 和 Kennedy 提出^[5]。其算法实际是一种基于迭代的优化工具, 每个优化问题的解都可以被想象成 d 维空间的一个点, 即称之为粒子, 每个粒子通过迭代搜寻, 在解空间追随两个最优的粒子来更新自己, 一个是粒子迄今为止寻找到的最优值, 叫做个体极值点($pbest$); 另外一个则是整个粒子群迄今为止寻找到的最优值, 叫做全局极值点($gbest$), 搜索这两个最优值后, 粒子通过如下两个公式更新自己, 直到找到食物位置。

$$v_i = w \cdot v_{i-1} + c_1 \cdot r_1 \cdot (pbest - x_i) + c_2 \cdot r_2 \cdot (gbest - x_i) \quad (1)$$

$$x_i = x_{i-1} + v_i \quad (2)$$

其中, v_i 是当前代的粒子移动速度, v_{i-1} 是前一代的粒子移动速度; r_1, r_2 是 0~1 间随机数, c_1, c_2 为学习因子, w 为惯性权重因子。

用 PSO 训练神经网络的具体步骤如下:

(1) 种群随机初始化。若构造的神经网络中需要确定的权值和阈值共 d 个, 则粒子群中每个粒子代表维解空间的一个点, 粒子的个数定为 K 。初始化每个粒子的速度 $v_i = (v_{i,1}, v_{i,2}, \dots, v_{i,d})$ 和位置 $x_i = (x_{i,1}, x_{i,2}, \dots, x_{i,d})$, 均取为 0~1 之间的随机数; 每个粒子的个体极值点 $pbest$ 初始值即为 x_i 的初始值。

(2) 计算每一个粒子的适应值。适应度函数采用如下误差函数

$$E = \frac{1}{P \cdot N} \sum_{p=1}^P \sum_{i=1}^N (y_i - \bar{y}_i)^2 \quad (3)$$

这里, P 代表了样本个数, N 代表输出单元的个数, \bar{y} 为期望输出。

(3) 根据适应值复制种群。先搜索全局极值点位置 $gbest$, 即为整个粒子群中取得最小个体极值的粒子的位置。若个体适应值 $E(x_i)$ 小于个体极值 $E(pbest)$, 则令个体极值位置 $pbest = x_i, E(pbest) = E(x_i)$ 。

(4) 粒子群位置更新与输出。根据公式(1)和(2)进行更新, 若 $E(gbest)$ 小于设定的控制值 ε 或迭代次数达到设定的进化次数 T 时, 停止更新, 搜索过程结束, 输出结果。反之转步骤(2)。

由于惯性权重因子 w 取值和最大速度 v_{max} 对程序运行结果有着很大的影响, 采用较大的 w 和 v_{max} 值能够快速搜索到粒子群体的全局目标区域, 但搜索精度不高; 而采用较小的 w 和 v_{max} 值可使种群的局部寻优能力增强, 搜索精度提高, 但其搜索速度较慢。因此, 兼顾各自的优点, 采用自适应的惯性权重因子进行搜索^[6,7], 即令

$$\omega = \omega_{max} - \frac{j(\omega_{max} - \omega_{min})}{T} \quad (4)$$

这里 j 为当前迭代次数, T 为设定的进化控制次数。

2.2 基于黄金分割的隐含层节点数的优化算法

神经网络的隐含层节点数 L 的最优选取是神经网络结构设计中最为重要的一个环节, 由于隐含层的节点数与网络的逼近精度和泛化能力息息相关, 因此在充分考虑其逼近精度的同时, 尽量使其泛化能力较强, 可以采用黄金分割优选法来确定合适的 L 值^[8]。

若输入层和输出层的节点数分别为 M 和 N , 理想的隐含层节点数 L 一般在如下范围内选取^[8]:

$$a = \frac{M+N}{2} \leq L \leq (M+N)+10 = b \quad (5)$$

基于黄金分割原理的具体优化算法步骤如下:

(1) 在 $[a, b]$ 上先确定第一个黄金分割点 $g_1 = 0.618 \times (b-a) + a$, 以 g_1 作为隐含层节点数, 求出网络输出误差 $E(g_1)$ 。

(2) 确定第二个黄金分割点, 即 g_1 的对称点 $g_2 = 0.382 \times (b-a) + a$, 求出网络输出误差 $E(g_2)$ 。

(3) 比较两误差大小, 舍劣取优。若 $E(g_1)$ 为优(即误差小), 则去掉 $[a, g_2]$, 留下 $[g_2, b]$; 若误差 $E(g_2)$ 小, 则去掉 $(g_1, b]$, 留下 $[a, g_1]$; 若两误差相等, 则去掉 $[a, g_2]$ 和 $(g_1, b]$, 留下 $[g_2, g_1]$ 。

(4) 在留下区间内类似上述过程, 即能快速搜索到理想的隐含层节点数 L 。并有如下关系:

$$E(L) = \min\{E(a), E(b), E(g_1), E(g_2), \dots\} \quad (6)$$

这里, g_1, g_2, \dots 为 $[a, b]$ 中的黄金分割试验点位置。

2.3 系统设计

图 1 为一种计算话务量的流程图。

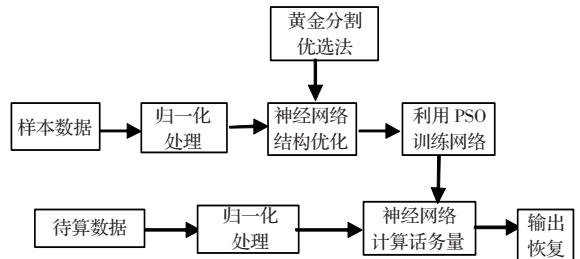


图 1 计算话务量的进化神经计算方法流程

(1) 样本信息预处理: 选取完备的样本信息, 并进行资料标准化、归一化处理等。

(2) 网络结构优化: 先依据样本信息的维数来设计网络的输入和输出层节点数, 再采用基于黄金分割的优化算法确定隐含层的节点数。

(3) 网络学习训练: 将归一化处理后的样本集输入神经网络进行学习训练, 直到满足精度要求或约定的循环次数为止, 存储权值和阈值。

(4) 计算与输出、恢复: 待计算的样本经过标准化、归一化处理输入训练好的神经网络进行计算并输出, 将输出结果按归一化的逆过程进行恢复数值。

3 实际应用与对比分析

采用河北省某市网通公司所提供的从 2004 年 11 月份到 2005 年 3 月份之间的 17 组关于话务量、来话接通率、掉话率和阻塞率之间关系小灵通业务数据进行实验, 其数据样本集如表 1 所示, 其中来话接通率 T 、掉话率 D 、无线阻塞率 B 作为神经网络输入, 周忙时话务量 a 作为网络的输出。所以神经网络的输入层有 3 个节点, 输出层有 1 个节点。

表1 样本集

时间	来话接通率/%	掉话率/%	无线阻塞率/%	周忙时话务量(erl)
11.13-11.19	47.09	2.07	1.920 0	1 838.021 4
11.20-11.26	47.58	2.15	3.349 5	2 010.204 4
11.27-12.03	48.12	2.14	3.125 2	2 151.772 4
12.04-12.10	47.69	2.01	3.575 7	2 095.092 8
12.11-12.17	48.44	2.06	3.475 6	2 123.571 3
12.18-12.24	48.07	2.13	3.671 9	2 150.971 3
12.25-12.31	47.89	2.01	3.754 2	2 153.213 3
01.01-01.07	47.51	2.16	4.600 7	2 264.493 5
01.08-01.14	46.90	2.37	2.486 9	1 685.644 9
01.15-01.21	46.73	2.32	1.900 9	1 728.842 8
01.22-01.28	47.99	2.16	2.911 3	2 045.825 4
01.31-02.06	48.51	2.25	4.998 1	2 359.325 6
02.07-02.11	48.70	1.49	4.469 0	1 839.406 0
02.12-02.18	46.31	2.20	2.910 0	1 851.848 0
02.19-02.25	47.89	2.01	3.217 6	2 032.178 9
02.26-03.04	47.62	1.99	1.190 0	2 103.041 8
03.07-03.11	48.44	2.23	3.706 1	2 142.340 3

3.1 训练过程

为保证学习拟合的精度和降低学习训练时间,选取表1中的前13组数据作为训练样本。

(1)归一化处理:采用如下归一化公式进行处理。

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (7)$$

这里 x 分别代表了三个输入接通率、掉话率及阻塞率,并且 $x \in [x_{\min}, x_{\max}]$, x_{\min}, x_{\max} 为最小值和最大值。三个输入的取值区间分别为: $[0.45, 0.50]$, $[0.01, 0.025]$, $[0.01, 0.05]$ 。

对网络的期望输出值采用如下归一化处理公式,其中 y 表示话务量。

$$y' = \frac{y}{10\ 000} \quad (8)$$

(2)学习训练:采用 PSO 算法,设粒子个数 $K=50$,最大速度 $v_{\max}=0.5$,最大惯性权重因子 $\omega_{\max}=0.9$,最小惯性权重因子 $\omega_{\min}=0.3$, r_1, r_2 取 0~1 之间的随机数, $c_1=2, c_2=2$,进化控制次数 $T=1\ 000$ 。

(3)隐含层节点数 L 的确定:采用基于黄金分割的优化算法,神经网络为“3-L-1”结构,则可以在 $2 \leq L \leq 14$ 之间范围内寻优,在各个黄金分割点中进行寻优, $E(L) = \min\{E(2), E(14), E(9), E(7), E(11), E(10)\}$,所以最终得到理想的 $L=9$,其网络输出误差为 $E(9)=2.489\ 3 \times 10^{-6}$,图2为其误差曲线图。

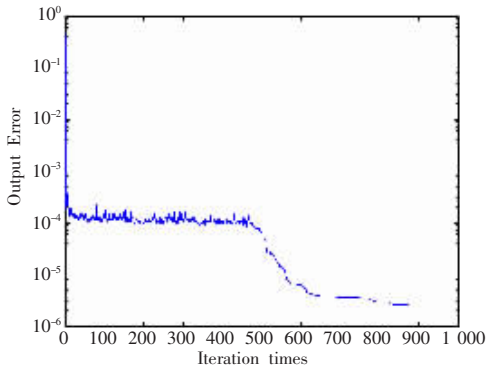


图2 PSO算法误差曲线图

(4)对比分析:为了检验 PSO 算法与其它算法的仿真效果,分别采用基本 BP 算法、改进 BP 算法进行学习训练。迭代

次数均取 1 000 次,误差曲线图分别如图3和图4所示,最终得到的训练结果如表2所示。

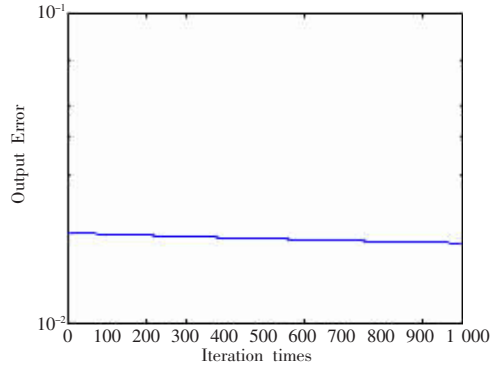


图3 基本BP算法误差曲线图

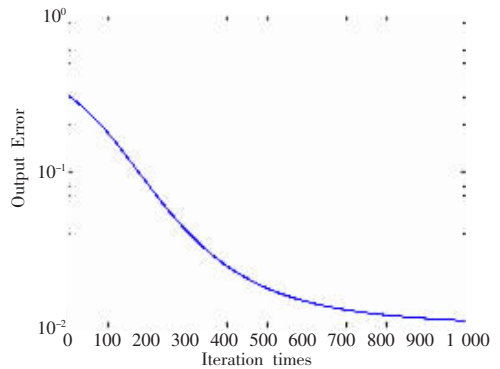


图4 改进BP算法误差曲线图

表2 训练结果

实际话务量	PSO	改进BP	BP
1 838	1 838	668.55	960.64
2 124	2 124	3 870.9	849.52
2 264	2 258	1 291.1	2 958.4
2 010	2 024	770.84	553.89
2 152	2 107	2 741.6	764.43
2 095	2 107	1 355.4	306.68
1 686	1 680	1 160.5	2 139.2
2 359	2 360	1 912.1	3 901.1
1 729	1 730	901.84	2 595.5
2 153	2 149	2 155	299.56
1 839	1 837	190.2	196.89
2 046	2 069	1 528.9	884.87
2 151	2 149	2 760.7	560.44
系统误差	2.489 3×10 ⁶	0.009 472 04	0.018 049 3

由图2至图4的误差曲线和表2结果可以看出,PSO 算法不仅收敛速度快,而且系统误差最小。

3.2 计算结果

选用表1中的后4组数据作为计算样本进行测试,神经网络的输出值经过归一化恢复后,最终计算结果如表3所示,其

表3 测试结果

实际话务量	PSO	BP	改进BP
2 142.340 3	2 144.5	404.21	615.94
2 103.041 8	2 080.3	8 264.5	157.62
1 851.848 0	1 852.7	4 955.0	5 721.8
2 032.178 9	2 022.2	2 556.1	2 860.6
方差	1.555×10 ²	1.272×10 ⁷	5.444×10 ⁶