

基于粒子群优化的过程神经网络学习算法

刘坤^{1,2} 谭莹^{1,2,†} 何新贵^{1,2}

1. 北京大学机器感知与智能教育部重点实验室,北京 100871; 2. 北京大学信息科学技术学院,北京 100871;

† 通信作者, E-mail: ytan@pku.edu.cn

摘要 基于粒子群优化为过程神经网络提出了一种新的学习算法。新算法在对网络输入函数和连接权函数进行正交基函数展开后,将网络中的结构参数和其他参数整合成一个粒子,再用粒子群优化算法进行全局优化。新算法不依赖于函数梯度信息,不需要手动调节网络结构。粒子群优化具有良好的全局优化性能和收敛性能,保证了过程神经网络的全局学习能力和新学习算法的收敛能力,更好地发挥过程神经网络的逼近性能。两个实际预测问题的实验结果表明,基于粒子群优化的学习算法比现有的基于梯度的基函数展开方法以及误差反传神经网络模型具有更好的预测精度。

关键词 过程神经网络; 学习算法; 粒子群优化; 基函数展开

中图分类号 TP183

Particle Swarm Optimization Based Learning Algorithm for Process Neural Networks

LIU Kun^{1,2}, TAN Ying^{1,2,†}, HE Xingui^{1,2}

1. Key Laboratory of Machine Perception (MOE), Peking University, Beijing 100871; 2. School of Electronics Engineering and Computer Science, Peking University, Beijing 100871; † Corresponding author, E-mail: ytan@pku.edu.cn

Abstract This paper proposes a new learning algorithm for process neural networks (PNNs) based on particle swarm optimization (PSO), called PSO-LM. After the orthogonal basis function expansion to the input functions and the weight functions of the PNN, the structure parameters and other parameters in the PNN will be formed as a particle, and globally optimized by PSO. This algorithm does not need any gradient calculations or the manual control of the network's structure. The global learning capability and the convergence capability of the PNN can be guaranteed by the capabilities of PSO, so the PSO-LM can better develop and improve the approximation capability of the PNN. According to two practical prediction applications, PSO-LM can outperform the existing basis function expansion based learning algorithm (BFE-LM) for PNNs, and the classic back propagation neural networks (BPNNs) on predictive accuracy.

Key words process neural networks; learning algorithm; particle swarm optimization; basis function expansion

过程神经网络^[1]针对许多系统中输入依赖于时间变化,输出既与输入信号的空间聚合作用有关,同时也依赖于时变输入信号的时间累积效应的情况而提出。过程神经元的输入和连接权值为时变函数(或函数向量),而输出为一个非时变值(或向量)。过程神经元在传统神经元空间聚合运算的基础上,增加了一个对于时间的累积算子,使过程神经

元同时具有对时、空二维信息的直接处理能力。由这类过程神经元构成的前馈过程神经网络模型的连续性、函数逼近能力和计算能力等理论问题均已解决。

粒子群优化算法^[2]是一种基于社会交互现象的随机全局优化技术,已经被成功地应用于训练人工神经网络等问题^[3-4]。粒子群优化算法易于实

国家自然科学基金(60875080, 60673020)和 863 计划(2007AA01Z453)资助

收稿日期: 2010-01-31; 修回日期: 2010-05-13; 网络出版日期: 2010-12-29

网络出版地址: <http://www.cnki.net/kcms/detail/11.2442.N.20101229.1122.006.html>

现,不依赖于梯度信息的计算,并且不易限于局部极值的特点使得它越来越多的被应用于各种优化问题中,例如函数优化、模糊系统控制、盲源分离以及机器学习^[5]。近些年来,对于粒子群算法的改进技术不断涌现,根据不同需求进行的有针对性的改进能够使得粒子群算法在具体应用中发挥出更好的性能。其中我们也对粒子群算法做了如下改进:引入克隆技术的克隆粒子群算法^[6],引入放大机制的放大粒子群算法^[7],以及受自然界中黑洞现象的启发而提出的随机黑洞粒子群算法^[8]等。这些算法都试图对粒子群中包含的信息量进行更充分地利用,以提高粒子的搜索效率,同时引入随机机制,保证种群的全局搜索能力,从而使得算法的性能得到全面提升。

本文基于粒子群优化的特点为过程神经网络提出了一种新的学习算法。对输入函数和连接权函数进行基函数展开后,模型中存在的参数如展开项数、隐层神经元个数和神经元阈值等都由粒子群优化算法进行调整和优化。这样,过程神经网络的隐层结构和参数可以被整体优化,而不需要使用试错法和基于梯度的方法。并且,粒子群算法的收敛性能和全局优化能力能够保证过程神经网络学习过程的有效性和现实性。

1 背景

1.1 过程神经网络

过程神经元的基本结构如图 1 所示。该过程神经元由时变输入信号加权,时间、空间二维聚合和阈值激励输出等 4 部分运算组成,其空间聚合算子取为多输入信号加权和,时间聚合算子为含时间变参积分。

图 1 中, $x_1(t)$, $x_2(t)$, ..., $x_n(t)$ 为过程神经元的连续时变输入函数; $w_1(t)$, $w_2(t)$, ..., $w_n(t)$ 为相应的连接权函数; $K(\cdot, \tau)$ 为过程神经元的聚合算子,空间上取多输入信号的加权和 \sum , 时间上取

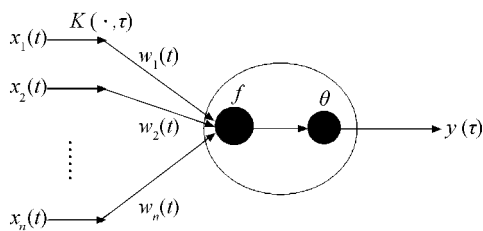


图 1 过程神经元模型
Fig. 1 Process neuron

含时间变参积分 \int_{τ} , f 为激励函数。

图 2 中展示了一个由过程神经元连接组成的单隐层前馈过程神经网络模型,其输入与输出之间的映射关系为

$$y = g \left(\sum_{j=1}^m v_j f_j \left(\int_0^T \left(\sum_{i=1}^n w_{ij}(t) x_i(t) \right) dt - \theta_j \right) - \theta \right), \quad (1)$$

其中, $[0, T]$ 为过程神经元输入信号的时间区间, θ_j 为第 j 个隐单元的阈值, θ 是输出元的阈值; $w_{ij}(t)$ ($i = 1, 2, \dots, n; j = 1, 2, \dots, m$) 为输入层神经元与隐层神经元的连接权函数; v_j ($j = 1, 2, \dots, m$) 为隐层神经元到输出神经元的连接权函数。 f 和 g 分别是隐层和输出层神经元的激励函数。

1.2 过程神经网络现有的学习算法

现有的过程神经网络学习算法主要包括基函数展开学习法^[9]和数字化学习法^[10]。

以输入数据为时间函数为例,基函数展开法分两步进行:首先,将输入函数和连接隐层的权函数在选定的正交基函数上进行有限项展开,利用标准正交基之间相乘变为常数 0 或 1 的性质,消去与时间有关的参数;然后,模型中的参数都使用传统的误差反传法进行调整,直到收敛。在第一步操作过后,模型中就不再包含与时间有关的参数,简单的梯度方法就可以应用进来,从而完成对模型参数的学习。然而,这种学习方法存在以下不足。1) 正交基的选取与展开项的个数都与具体的问题有关系。其中展开项个数 L 的确定对网络十分重要,少了可能丢失原函数中的信息,导致学习的知识不完善;多了则可能引入原信号中的噪音信号使得学习的知识不准确。基函数展开操作后, L 的取值还将影响隐层神经元的结构,因此间接影响了网络的结构。在基函数展开法中 L 的确定往往是凭经验和试错法来进行的。这种方法效率低,而且难以做出最优的决定,从而会影响模型的性能。2) 网络隐层的神经元点个数也同样需要凭经验和试错法来确定,很难达到最优。3) 梯度下降的方法存在固有的容易早熟,以及需要

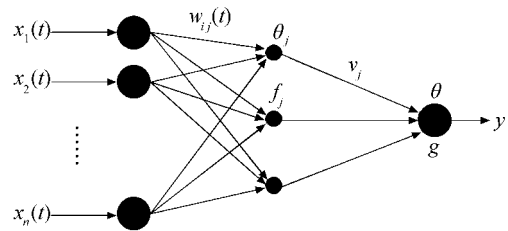


图 2 前馈过程神经网络模型
Fig. 2 Feedforward process neural network

计算复杂的梯度信息等缺点。总之,传统的基函数展开方法既没有对网络的结构起到优化的功能,又在参数优化中存在梯度下降法固有的缺陷。

数字化算法将离散时间函数作为输入,权函数也要做相应的离散化,与输入的各分量对等。例如,输入信号长度为 T ,则连接它与下一层的每个权函数中都会有 T 个参数来与之对应。假设隐层神经元个数为 m ,输入神经元个数为 n ,则这一层权函数的参数个数就为 $T \times m \times n$ 个。这样会导致如下问题:首先,如果输入序列长度过大,会导致需要学习的参数非常多,计算量迅速增加,造成学习困难;其次,数字化方法对于权函数的设计过于细致,容易造成过拟合,导致模型的泛化能力不足;此外,这种方法也依赖于误差反传算法,进而受制于梯度计算和容易早熟的问题。特别当有些函数的梯度计算比较复杂,甚至不可计算时,也会导致算法出现计算问题。

总之,为过程神经网络寻找一种既能优化其网络结构,又能学习其网络参数,既能进行全局性的优化,又能控制计算复杂度的学习算法是本文研究的目的。

1.3 粒子群优化算法

Bratton 等^[11]定义了一种标准的粒子群优化算法。粒子更新公式表达如下:

$$V_{id}(t+1) = \chi(V_{id}(t) + c_1 \varepsilon_1(P_{iBd}(t) - X_{id}(t)) + c_2 \varepsilon_2(P_{gBd}(t) - X_{id}(t))), \quad (2)$$

$$X_{id}(t+1) = X_{id}(t) + V_{id}(t+1), \quad (3)$$

其中, $i = 1, 2, \dots, n$, n 是种群中粒子的个数, $d = 1, 2, \dots, D$, D 是解空间的维数。 c_1 和 c_2 叫做学习因子,是非负常数, r_1 和 r_2 是 $[0, 1]$ 区间中的均匀随机数,对于每个粒子的每一维上都有不同的取值。 P_{iBd} 和 P_{gBd} 分别是第 i 个粒子和其邻域中所有粒子在第 d 维上的所找到的最优位置。收缩系数 χ 如下定义:

$$\chi = \frac{2}{|2 - \varphi - \sqrt{\varphi^2 - 4\varphi}|}, \varphi = c_1 + c_2. \quad (4)$$

Bratton 等^[11]同时建议其中的参数设置为 $c_1 = c_2 = 2.05$, $\chi = 0.72984$ 。算法的停止准则为适应度水平达到指定的要求,或者适应度评价次数达到上限。

2 基于粒子群优化的过程神经网络学习算法

2.1 学习算法的提出

我们考察图 2 中所示的前馈过程神经网络,隐层中所有 m 个神经元都是过程神经元,其输出为

常数;输出层只有一个普通神经元。其中权函数只出现在输入层到隐层之间,其他连接权及所有阈值都是常数。在这个过程神经网络的学习过程中,最重要的就是学习连接权函数,由于函数空间是无穷维的,无法用穷举的方法直接学到,所以仍然使用基函数展开的方法来解决这个问题。基函数的选取需要根据具体问题来确定,但是基展开的项数 L 留给粒子群算法来决定。另外,隐层的神经元个数 m 以及模型中的其他参数都由粒子群算法统一进行优化。

考察基函数展开后式(1)所示的模型输出的变化:

$$y = g\left(\sum_{j=1}^m v_j f_j\left(\sum_{i=1}^n \sum_{l=1}^L a_{il} w_{ij}^{(l)} - \theta_j\right) - \theta\right), \quad (5)$$

其中 a_{il} 是第 i 个输入信号基函数展开后的第 l 项系数,是已知的常数。而式中未知的需要学习和调整的参数有展开项数 L 、隐层神经元数 m 、隐层连接权函数的展开系数 $w_{ij}^{(l)}$ 、隐层神经元阈值 θ_j ($j = 1, 2, \dots, m$)、输出神经元阈值 θ 及其连接权 v_j ($j = 1, 2, \dots, m$)。这些参数中既有关于网络结构的(隐层神经元个数 m),也有关于基函数展开操作的(展开项数 L),还有关于权值和阈值的其他参数。各种参数的取值范围也有很大差异: m 和 L 的取值为正整数,而其他参数的取值为实数。

基于粒子群优化的特点,把上述参数统一放到 PSO 模型中进行整体优化。这样可以使得针对训练目标的过程神经网络的学习更具有整体性,并充分利用粒子群算法的全局搜索能力,使得 PNN 能更好的发挥其强大的逼近能力。PSO 中一个粒子可以表示为

$$p = (L, m, w_{ij}^{(l)}, \theta_j, v_j, \theta), i = 1, \dots, n, j = 1, \dots, m, l = 1, \dots, L, \quad (6)$$

这样粒子的维度为

$$d = 1 + 1 + nmL + m + m + 1 = m(nL + 2) + 3. \quad (7)$$

粒子群算法中的主要参数(如种群中粒子的个数、惯性权值、学习因子、拓扑结构以及停止准则等)的取值可以参考文献[11]。本算法中需要重点考虑的设置则是粒子的搜索空间的合理限定,这是充分发挥粒子群同时优化 PNN 结构和参数能力的保证。

首先,从粒子的构成式(6)来看,前两维的取值范围为正整数,并且不能过大。 L 的取值范围还受

到具体的基函数选取的影响,例如小波基展开的最大项数根据小波基的阶数不同是不一样的。隐层神经元的个数 m 需要根据问题的复杂程度设定一个上限。其他维上的取值范围应为连续的实数区间,可以根据具体问题的目标值所在范围予以合理限定。

另外根据式(7),粒子维度的大小即搜索空间的维度大小,主要受前两维取值的影响。如果 L 和 m 的取值范围过大,则可能导致粒子维度的过高;而如果取值范围限制过紧,则可能导致网络的结构过于简单从而没有足够的逼近能力,进而导致学习能力的瓶颈。总之,上述参数的限定要根据具体的基函数和具体的问题背景来确定。在实验中,我们将结合问题的特点来决定粒子搜索的范围。

对于算法的停止准则,可以使用经典的交叉验证方法。合理的限定 PSO 的迭代步数以防止发生对训练样本的过拟合,保证学习算法的泛化能力。具体来说,在 PSO 的迭代过程中,当在交叉验证法的确认子集上的预测精度达到最低时,可以停止学习。另外,在学习过程中使用正则项方法设置早停规则也是一种有效的方法。

综上所述,基于粒子群优化的过程神经网络学习算法如下。

步骤 1 将过程神经网络训练的性能函数设置为粒子群算法的目标函数,并选择正交函数基。

步骤 2 将过程神经网络中的参数按式(6)顺序编号成为种群中粒子的各维。

步骤 3 设置粒子群算法中主要参数,限定粒子搜索空间,确定算法停止准则。

步骤 4 运行粒子群算法直至停止准则满足。

步骤 5 如果结果不满意,回到步骤 3。

2.2 学习算法的讨论

从上述算法(下简称为 PSO-LM)的设计中可以看出,算法的核心在于 PSO 粒子的设计和粒子搜索空间的限定。需要被优化的量都可以按顺序放入粒子中,作为粒子的一维。其中展开项 L 和隐层神经元个数 m 的加入,是本算法的一项重要设计。PSO 将这些因素从整体上进行优化,不但解决了展开项如何确定的问题,也解决了网络隐层结构的确定问题,而且使得结构和参数的学习更有整体性和全局性。

PSO-LM 克服了我们先前分析的已有算法的缺点,并且拥有一些新的特点。首先,PSO-LM 设计简

单,PSO 的引入和实现非常直接,本身的搜索策略并不复杂。其次,过程神经网络的结构参数 L 和 m 由 PSO 进行优化,不必进行手动调整,使得算法更能从整体上自动对模型性能进行优化。最后,PSO 本身具有极强的全局优化能力和收敛性能,既避免了使用梯度法导致的计算困难和早熟现象,也能保证算法的学习能力和收敛性能。PSO 从全局出发对于目标映射进行的整体学习可以提升 PNN 的逼近性能,只要训练样本的代表性可以保证,则新算法就能更准确的学习到映射的真实特征,从而保证在预测和控制问题中的泛化性能。

PSO-LM1 的计算复杂度主要由种群中粒子个数以及 PSO 迭代次数决定。PSO-LM1 的一次迭代中需要为每一个粒子都进行性能评价(即计算一次网络前馈输出值)。假设种群中有 p 个粒子,则 PSO-LM1 是同时在对 p 个网络模型进行学习和调整,并且这些网络模型间存在交流和学习。相应的,BP 算法中每次迭代只调整一个模型,但是每次调整的策略是“贪心”的,并且是以计算梯度信息为代价的。所以,从理论上来说,两种算法的计算量各有优劣。在实际应用 PSO-LM1 时,可以通过合理的设定粒子个数和迭代次数来限制算法的计算量,利用 PSO 协同工作和全局优化能力强特点,从性能和计算量之间找到一个较好的权衡。最后,PSO-LM1 在解决相对复杂的问题上,尤其是梯度信息难以求解或者不能求解的大规模复杂问题上会有更大的优势。表 1 对 PSO-LM 与已有模型和算法的优劣进行了总结。

表 1 新算法与已有模型和算法的比较
Table 1 Comparisons among new algorithm and existing algorithms

模型算法	优点	缺点
BPNN	学习策略简单	隐层结构需要经验调节; 计算梯度,容易早熟; 对过程性信息学习效率低
BFE-LM	可以学习过程性权函数;基函数展开计算简单	基函数展开存在众多经验参数;计算梯度,容易早熟
PSO-LM	整体优化网络结构和权函数;全局优化能力强;没有复杂计算	粒子搜索空间需要慎重确定

3 实验

3.1 短期交通流预测问题

交通流预测问题,即根据已有的道路车流量数据预测道路某处下一时刻的车流量,短期预测问题通常为预测后面 5 分钟到半小时内的流量。数据来自于美国加州大学高速公路交通评价系统 PeMS。我们截取了其中第 7 区的连续 5 个线圈在 2009 年 6 月的 21 个工作日内从上午 10:00 到下午 15:00 的数据,每 5 分钟记录一次。

实验的目标是使用全部 5 个线圈当前的数据来预测中间线圈未来 5 分钟的车流量^[10],比较已有的普通过程神经网络基展开算法(BFE-LM),传统的误差反传神经网络(BPNN)以及新的基于粒子群优化的过程神经网络学习算法(PSO-LM)的预测性能。前两周的数据(每天包含 110 个样本)用来训练已有模型,第 3 周的数据(每天包含 55 个样本)用来测试模型预测精度。3 个模型的主要参数设置见表 2。

表 2 中的 BPNN 模型为单隐层前向人工神经网络,隐层神经元 25 个;BFE-LM 模型为单隐层的前向过程神经网络,隐层神经元 25 个,隐层基函数取 5 阶 db1 小波基,展开项取前 10 项;PSO-LM 模型为单隐层前馈过程神经网络,隐层基函数取 5 阶 db1 小波基,展开项的项数 L 以及隐神经元个数 m 均由 PSO 优化决定。另外,经过试验 3 个模型的激励函数均应设为线性函数 purelin。

表 2 交通流预测实验模型参数设置

Table 2 Parameter setups of traffic flow prediction experiment

模型	网络结构	权函数操作
BPNN	5-25-1	常数
BFE-LM	5-25-1	5 阶 db1 小波基 $L=10$
PSO-LM	5- m -1	5 阶 db1 小波基

在 PSO-LM 模型中,其参数的设置还包括学习因子 $c_1 = c_2 = 2.05$,收缩系数 $\chi = 0.72984$,种群粒子个数 $p = 100$ 。此外,粒子搜索空间的限制如下: L 的取值为正整数,下限为 1,上限为小波基展开后的最大项数; m 的取值为区间 $[1, 15]$ 内的整数;其他参数的取值范围为 $[-2, 2]$ 。

预测交通流量的性能指标除了每次预测平均消耗的时间外,还有如下两个常用误差指标:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}, \quad (8)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i}, \quad (9)$$

其中式(8)为均方根误差,单位为车辆数;式(9)为平均绝对百分比误差,评价的是预测误差与真实流量的相对值,没有单位。 n 为时间取点个数,即样本个数; \hat{y}_i 为模型预测出的 t 时刻的结果,而 y_i 为 t 时刻的真实结果。我们对 3 个模型在工作日的 5 天中的预测结果分别进行比较,每个模型将独立运行 30 次,记录其平均的性能结果。实验的统计结果如图 3 所示。

图 3(a)和(b)分别为 3 个模型的预测精度比较图,横坐标为一周的 5 天,纵坐标为预测的误差指标;(c)展示了模型对数据进行一次训练和预测的平均耗时。从预测精度上看,使用 PSO-LM 的过程神经网络有着最好的性能,除了在极个别情况下的预测精度(周五)略逊于 BFE-LM,在所有实验中都好于 BPNN;BFE-LM 训练的过程神经网络在周五的预测中最准确,但是整体性能起伏较大,并且在大多数情况下都难以得到最好解,因为梯度下降法的早熟问题导致结果不稳定;BPNN 虽然在某些情况下性能接近 PSO-LM,但是也是由于梯度下

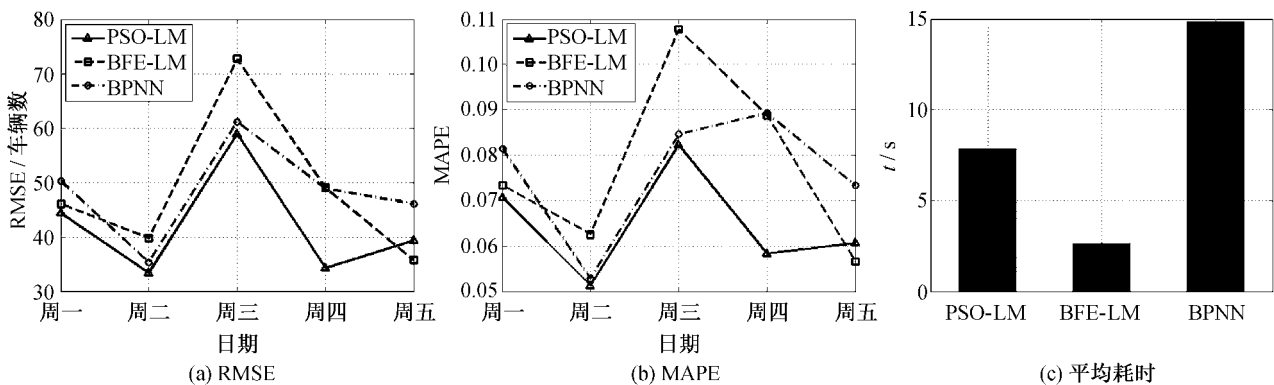


图 3 交通流预测实验统计结果

Fig. 3 Statistic results of traffic flow prediction experiment

降法的早熟等问题而无法稳定的提供准确的预测。从预测耗时上来看, BPNN 耗费的时间也最多, 因为它对于输入信号的所有内容都直接进行处理, 当信号长度较大时, 时间消耗大的问题将会更加突出; 而其他两种模型都用了基函数展开, 对原始数据进行了主成分的抽取, 将样本的信息集中于少数主成分上, 为后面的处理提供了便利, 所以用时较短; PSO-LM 所消耗的时间比 BFE-LM 多, 是因为 PSO 中每一次迭代要对所有粒子都进行一次网络前馈取值, 来评价粒子的适应度值。但是多个粒子的相互协作交流能够更快更有效的找到最优点, 从而能够降低搜索的步数, 所以整体时间消耗仍是可以接受的。

分别对预测误差指标 MAPE 和 RMSE 做方差分析, 检验 3 个模型的预测结果的差异是否具有统计意义。分析得到 p 值分别为 0.0255 和 0.0016, 分别小于 0.05 和 0.01, 表明新算法与旧算法在预测性能上的差异分别具有统计意义和高度统计意义, 即新算法对 PNN 性能的提高是有统计意义的。这说明 PSO-LM 使得 PNN 获得了更好的学习能力和泛化能力, 从而可以对未见样本做出精度更高的预测。此外, 由 PSO-LM 确定的 PNN 中权函数展开项数 $L = 6$, 隐层神经元个数 $m = 13$ 。

3.2 地壳形变数据预测问题

为了再次比较上述实验中 3 个模型对时间序列的拟合和预测能力, 我们使用国家地震科学数据共享中心提供的某地区地壳形变南北向伸缩数据^[12], 进行地壳形变序列预测实验。数据为某地震台记录的 2008 年当地地壳形变南北向的伸缩值, 每小时记录一次。实验的目标是根据现有的数据预测未来 24、48 和 72 小时的伸缩值, 预测精度仍用式 (8) 和 (9) 来衡量, 而平均每次训练和预测的耗时将用来显示各模型的计算量。

我们用前面 10 组已知数据来预测未来的值, 例如, 要预测未来第 24 小时的值, 则需要前面 10 个相互间隔 24 小时的已知数据作为一个样本。这样, 我们取了 1000 个训练样本用于训练模型, 另外 1000 个测试样本用于检验模型的预测性能。各个模型的参数设置如表 3 所示。

3 个模型的激励函数均为线性函数 purelin。另外, PSO-LM 中粒子每一维度上的搜索区域定于 $[-2, 2]$, 粒子个数设为 150, L 和 m 的取值范围分别为 $[1, 10]$ 和 $[1, 25]$ 上的整数。其他参数均与实

表 3 地壳形变预测实验模型参数设置
Table 3 Parameter setups of crustal deformation prediction experiment

模型	网络结构	权函数操作
BPNN	10-25-1	常数
BFE-LM	1-25-1	5 阶 db1 小波基, $L = 10$
PSO-LM	1- m -1	5 阶 db1 小波基

验 1 相同。各个模型均经过 30 次独立运行, 实验平均结果如表 4 所示。

从表 4 中可以看出, PSO-LM 的预测精度明显高于其他两个模型, 说明 PSO-LM 使得 PNN 获得了更好的学习能力, 并得到了更好的泛化性能。PSO-LM 耗时较多, 原因在于每次迭代都要对 150 个粒子进行适应度值评价。

对比 BFE-LM 和 BPNN 发现, 两个算法的预测精度和耗时都相差不大, 证明单纯使用 BFE 的 PNN 相对于 BPNN 没有优势。然而同时使用 BFE 和 PSO 的 PSO-LM 则对预测精度有明显的提高, 说明将 PSO 引入 PNN 的学习算法是有效的。同时, 这个实验结果表明 PNN 的固有逼近能力是强于 BPNN 的, 但是如果学习方法不当则可能使得 PNN 的逼近能力无法完全发挥。而我们所提出的 PSO-LM 则能够发挥出 PNN 的潜力, 使得 PNN 在具体应用中的性能有明显提升。虽然 PSO-LM 的计算时间较多, 但是考虑到现代计算机的计算速度不断加快, 新算法多一些时间的花费是可以接受的。

进一步分析实验结果可以看出, 问题越复杂, PSO-LM 对原有算法的性能提高越明显。一方面, 前两种算法在面对复杂问题时, 梯度计算更难且更容易陷入早熟困境。另一方面, PSO-LM 中粒子群

表 4 地壳形变预测实验统计结果
Table 4 Statistic results of crustal deformation prediction experiment

模型	预测步数	RMSE	MAPE	t/s
BPNN	24	6.5597	0.0122	1.13
	48	6.3414	0.0112	1.20
	72	20.4035	0.0379	1.83
BFE-LM	24	6.5597	0.0122	1.29
	48	6.3414	0.0112	1.31
	72	18.5977	0.0346	2.08
PSO-LM	24	5.2323 [*]	0.0091	8.03
	48	5.5387 [*]	0.0094	10.76
	72	7.8416 [*]	0.0137 [*]	14.04

注: * 表示提高量在方差分析 5% 水平上具有统计显著性。

互相合作交流获取信息的优势在复杂问题上更能发挥出来。这说明使用 PSO-LM 的 PNN 在解决复杂的实际问题中将会有更大的潜力。

4 结束语

本文利用粒子群算法便于实现、全局优化能力强、不需要梯度计算等优势,为过程神经网络提出了一种新的基于粒子群优化的学习算法(PSO-LM)。本算法可以优化过程神经网络的结构参数,包括其基函数展开项数以及隐层神经元个数和其他常数参数,粒子群算法的全局搜索能力和收敛能力保证了算法的实效性。实验结果表明新算法使得 PNN 具有更好的学习能力和泛化能力,而计算量的增加仍可接受。需要注意的是,合理确定 PSO 中粒子各维度的搜索范围是 PSO-LM 需要着重解决好的一个问题。显然,搜索范围确定需要根据具体问题的先验知识进行合理分析和具体调整,否则将制约 PSO-LM 在实际中的应用。

总之,PSO-LM 丰富了 PNN 的学习算法,为 PNN 更好的发挥逼近能力,更广泛的实际应用提供了有力的工具。

参考文献

- [1] 何新贵,许少华. 过程神经网络. 北京: 科学出版社, 2007
- [2] Kennedy J, Eberhart R C. Particle swarm optimization // Proceedings of the IEEE International Conference on Neural Networks. Perth, Western Australia, 1995: 1942-1948
- [3] Ismail A, Engelbrecht A P. Training product unit in feedforward neural networks using particle swarm optimization // Proceedings of the International Conference on Artificial Intelligence. Stockholm, 1999: 36-40
- [4] Bergh V D F, Engelbrecht A P. Cooperative learning in neural networks using particle swarm optimizers. South African Computer Journal, 2000, 26: 84-90
- [5] Eberhart R C, Shi Yuhui. Particle swarm optimization: developments, applications and resources // Proceedings of the IEEE Congress on Evolutionary Computation. Seoul, 2001, 1: 81-86
- [6] Tan Ying, Xiao Zhongmin. Clonal particle swarm optimization and its applications // IEEE Congress on Evolutionary Computation, Singapore, 2007: 2303-2309
- [7] Zhang Junqi, Liu Kun, Tan Ying, et al. Magnifier particle swarm optimization for numerical optimization // Proceedings of Genetic and Evolutionary Computation Conference. Atlanta, 2008: 167-168
- [8] Zhang Junqi, Liu Kun, Tan Ying, et al. Random black hole particle swarm optimization and its application // International Conference on Neural Networks and Signal Processing. Zhenjiang, 2008: 359-365
- [9] 许少华,何新贵. 基于函数正交基展开的过程神经网络学习算法. 计算机学报, 2004, 27(5): 645-650
- [10] Wu Tianshu, Xie Kunqing, Song Guojie, et al. Numerical learning method for process neural network // International Symposium on Neural Networks. Wuhan, 2009: 670-678
- [11] Bratton D, Kennedy J. Defining a standard for particle swarm optimization // Proceedings of the IEEE Swarm Intelligence Symposium. Hawaii, 2007: 120-127
- [12] 国家地震科学数据共享中心. 地形变数据库 [DB/OL]. (2002-01-01) [2010-03-20]. http://data-prc. enc. earthquake. cn/qzsj/d_mtd_060. htm