信息与控制　2018 年　第 47 卷　第 5 期：553 ~ 558
DOI：10.13976/j.cnki.xk.2018.7188　文章编号：1002 - 0411(2018) - 05 - 0553 - 06

一种增量极限过程神经网络的研究及应用

杜　娟[1]，刘志刚[1]，许少华[2]，高雅田[1]
1. 东北石油大学计算机与信息技术学院，黑龙江 大庆　163318；
2. 山东科技大学计算机科学与工程学院，山东 青岛　266590

基金项目：国家自然科学基金资助项目(61170132，41330313)；黑龙江省自然科学基金资助项目(F2015021)
通信作者：刘志刚，dqpilzg@163.com
收稿/录用/修回：2017 - 04 - 17/2017 - 07 - 19/2017 - 09 - 06

摘要

为提高过程神经网络逼近效率，从模型结构角度出发，提出了一种增量极限过程神经网络模型，根据输出误差在隐层中逐次加入新节点实现结构自增长. 首先利用量子衍生萤火虫算法优化新增临时节点输入参数；其次根据新增节点输出正交向量的 2 范数判别相关性；最后固定现有节点参数，通过极限学习理论计算新增节点的输出权值. 在仿真实验中，通过与其它过程神经网络对比分析，以 Henon 时间序列预测和页岩的岩性识别为例验证所提方法的有效性，模型逼近效率和训练速度均有提高.

关键词
极限学习
过程神经网络
正交向量
结构自增长
岩性识别
中图法分类号： TP183
文献标识码： A

Research and Application of Incremental Extreme Process Neural Network

DU Juan[1], LIU Zhigang[1], XU Shaohua[2], GAO Yatian[1]
1. *School of Computer and Information Technology*, *Northeast Petroleum University*, *Daqing* 163318, *China*;
2. *College of Computer Science and Engineering*, *Shandong University of Science and Technology*, *Qingdao* 266590, *China*

Abstract

To improve the approximation efficiency of process neural networks, we propose an incremental extreme process neural network from the model structure perspective, which realizes the adaptive growth of the hidden-layer structure by gradually adding new neurons to the hidden layer based on the output error. First, we propose a quantum-inspired firefly algorithm to optimize the input parameter of a newly added neuron. Then, we analyze the relevance between the new neuron and the existing neurons according to the orthogonal vector 2-norm with respect to the new neuron′s ouput. Lastly, we calculate the output weights of the new neuron based on the extreme learning theory while fixing the weight parameters of the existing neurons. Through a simulation experiment based on Henon time series forecasting and shale lithology identification, we compare the performance of the proposed method with those of other process neuron networks, and verify the effectiveness of our proposed method and the obvious improvements realized by the model′s approximation efficiency and training speed.

Keywords
extreme learning;
process neural network;
orthogonal vector;
structure self-growth;
lithology recognition

0　引言

2000 年，何新贵院士等提出过程神经网络(process neural network，PNN)[1]，利用时空二维聚合机制提高非线性时变系统的逼近效率. 为简化时域计算，许少华教授提出正交基展开输入和权函数的方法[2]，使 PNN 理论研究快速发展. 在模型研究方面，相继有反馈[3]、模糊[4]、离散[5]等模型提出；在学习算法方面，主要有梯度下降[2-4]和进化算法优化[6]两种. 目前 PNN 研究在以下方面仍存在一定的不足：

1）由于 PNN 映射机制复杂且梯度下降法对初值敏感，训练收敛性不高，而且进化算法优化等价于对多个 PNN 网络同时学习，因此计算非常耗时；

2）确定隐层规模时主要采用试探法，自适应结构的研究还很少，但隐层结构对模型的逼近效率有重要影响. 为此，刘坤等[7]在 2011 年提出使用粒子群算法优化 PNN 结构，取得一定效果. 但正交基展开后 PNN 参数成倍增长，优化算法面对高维问题极易陷入局部最小，严重影响网络逼近效率. 以上两个问题成为 PNN 实际应用中的瓶颈问题.

2004 年，新加坡南洋理工大学 Huang 提出极限学习机(extreme learning machine，ELM)理论[8]，通过理论分析和实验已证明可极大提高神经网络的学习速度[9]. 截至目前

以 ELM 为研究主题的文献共被 SCI 收录 2 606 篇，其中 372 篇为 IEEE 文献，在模式识别[10-11]、回归预测[12-14]、智能芯片[15-16]等领域都已成为国际研究热点. 2016 年，作者所在课题组根据 PNN 自身特性，融合 ELM 分别提出固定型极限过程神经网络(fixed ELM PNN，FE-PNN)[17]、进化型极限离散过程神经网络(QPSO ELM discrete PNN，QE-DPNN)[18]，较大提高了 PNN 学习速度.

本文在前期研究工作的基础上，针对 PNN 的隐层规模及训练算法问题，提出一种增量极限过程神经网络(incremental extreme PNN，IE-PNN)，采取逐次添加新节点直至满足输出误差的方式实现隐层结构自增长. 对于新增过程神经元节点：

1）输入到该节点的输入参数(正交基展开系数)使用量子衍生萤火虫算法优化计算；

2）为消除冗余节点，提出使用 Gram-Schmidt 正交化完成新增节点与现有节点的相关性分析；

3）根据极限学习理论，固定已有节点的权值系数，根据网络输出误差计算新增节点输出参数. 仿真实验通过 Henon 时间序列预测验证方法有效性，并实际应用于页岩开发的岩性识别问题中. 实验表明结构自增长机制不但可节省人为试探的时间，同时紧凑的模型结构可有效提升过程神经网络的逼近效率且具有较快的学习速度.

1 过程神经网络 PNN

过程神经元包括时空聚合两个算子，并且输入和权值均是与时间有关的函数，具体结构如图 1 所示. 其中，$x_i(t)$ 和 y 为模型输入与输出，f 和 g 为神经元激励函数，$w_{ij}(t)$ 为隐层输入权值函数，v_j 为隐层输出权值，θ_j 为过程神经元阈值，则输入输出的映射关系为

$$y = g\left(\sum_{j=1}^{m} v_j f\left(\int_0^T \sum_{i=1}^{n} w_{ij}(t)\cdot x_i(t)\,\mathrm{d}t - \theta_j\right)\right) \tag{1}$$

其中，T 为采样周期. 对输入 $x_i(t)$ 和权函数 $w_{ij}(t)$ 实施正交基函数 $b_l(t)$ 有限项展开：

$$\begin{cases} x_i(t) = \sum_{l=1}^{L} a_{i,l} b_l(t) \\ w_{ij}(t) = \sum_{l=1}^{L} w_{ij,l} b_l(t) \end{cases} \tag{2}$$

其中，$a_{i,l}$、$w_{ij,l}$ 为基展开系数，L 为满足拟合精度时的基函数展开项数.

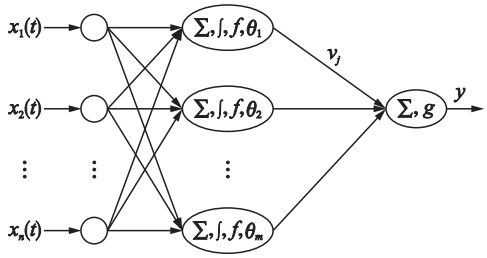

图 1 过程神经元网络

Fig.1 Process neural network

根据基函数正交特性，式(1)可以整理为

$$y = g\left(\sum_{j=1}^{m} v_j f\left(\sum_{i=1}^{n}\sum_{l=1}^{L} w_{ij,l} a_{i,l} - \theta_j\right)\right) \tag{3}$$

对于学习样本集 $\{X_k(t), d_k\}_{k=1}^{K}$，误差函数定义为

$$\begin{aligned} E &= \sum_{k=1}^{K}(y_k - d_k)^2 \\ &= \sum_{k=1}^{K}\left(\sum_{j=1}^{m} v_j f\left(\sum_{i=1}^{n}\sum_{l=1}^{L} w_{ij,l} a_{i,l}^{(k)} - \theta_j\right) - d_k\right)^2 \end{aligned} \tag{4}$$

其中，$a_{i,l}^{(k)}$ 为第 k 个样本的基展开系数.

2 增量极限过程神经网络(IE-PNN)

2.1 新增节点的输入权值优化计算

为方便描述，记 IE-PNN 模型有 n 个输入，隐层中现有过程神经元节点 m 个，新增节点序号为 $m+1$. 增量 ELM 理论对新增节点输入权值采取多次随机方式，利用输出误差下降速率确定新增节点输入参数. 但 PNN 模型复杂，新增节点输入权值为权函数基展开系数共 $n\times L$ 项，相当于增量 ELM 新增节点参数的 L 倍，此时随机赋值难获取到最优的模型参数. 本文根据群智能优化策略，对 IE-PNN 新增节点正交基展开系数实施全局寻优. 优化算法选取 2008 年英国剑桥大学学者 Yang 提出的萤火虫算法(firefly algorithm，FA)[19]. 该算法主要根据萤火虫的发光亮度引发个体间的相互吸引完成问题寻优，算法参数较少、易于使用，目前已有研究表明该算法相对于遗传算法 GA 和粒子群算法 PSO 优化效率更高[20]. 本文为进一步加快 FA 算法的搜索效率，提出一种量子衍生萤火虫算法(quantum-inspired FA，QFA)，关键算子为：

1）萤火虫的量子编码. 首先对新增节点的输入参数实施单位圆量子编码，第 i 只萤火虫的编码结构为

$$\boldsymbol{p}_i = \left[\begin{vmatrix}\cos\theta_{i1}\\ \sin\theta_{i1}\end{vmatrix}\begin{vmatrix}\cos\theta_{i2}\\ \sin\theta_{i2}\end{vmatrix}\begin{vmatrix}\cdots\\ \cdots\end{vmatrix}\begin{vmatrix}\cos\theta_{iD}\\ \sin\theta_{iD}\end{vmatrix}\right] \tag{5}$$

其中，问题维数 $D=n\times L$，量子幅角 $\theta_{ij}=2\pi\times\mathrm{rand}(0,1)$.

2）萤火虫的相对吸引力. QFA 算法优化新增节点输入参数过程中，萤火虫通过比较相对吸引力的大小确定自身的移动方向. 引入量子编码后，萤火虫 i 相对于萤火虫 j 的距离为 r_{ij} 二者在单位圆上的弧长，它们间的相对吸引力为

$$\begin{aligned} \beta_{ij} &= \exp(-\gamma\cdot r_{ij}^2) \\ &= \exp\left(-\gamma\left(\sum_{d=1}^{D}\frac{|\theta_{i,d}-\theta_{j,d}|\pi}{180}\right)^2\right) \end{aligned} \tag{6}$$

其中，γ 为光吸收系数.

3）萤火虫的位置更新. 萤火虫 j 向萤火虫 i 移动过程中，其量子位幅角的更新方式可描述为

$$\Delta\theta_j(t+1) = \beta_{ij}(\theta_i(t) - \theta_j(t))$$

此时根据量子旋转门，萤火虫 j 的每一个量子位概率幅的更新方式为

$$\begin{bmatrix}\cos(\theta_{ij}(t+1))\\ \sin(\theta_{ij}(t+1))\end{bmatrix} = \begin{bmatrix}\cos(\theta_{ij}(t)+\Delta\theta_{ij}(t+1))\\ \sin(\theta_{ij}(t)+\Delta\theta_{ij}(t+1))\end{bmatrix} \tag{7}$$

2.2 新增节点的相关性判别

与普通神经网络类似，PNN 模型结构最优时，每个隐层过程神经元节点对训练样本的输出应当线性无关，即 PNN 模型输出应为隐层输出向量的线性组合. 为此，本文

提出一种利用 Gram-Schmidt 正交化的方法完成新增节点的相关性判别. 按式(3)，隐层中第 j 个过程神经元对样本的输出向量 $\boldsymbol{h}_j$ 为

$$\boldsymbol{h}_j = [f(a_{i,l}^{(1)}, w_{ij,l}, \theta_j), \cdots, f(a_{i,l}^{(K)}, w_{ij,l}, \theta_j)] \tag{8}$$

其中，

$$f(a_{i,l}^{(k)}, w_{ij,l}, \theta_{\mathrm{m}}) = f\Big(\sum_{i=1}^{n}\sum_{l=1}^{L} a_{i,l}^{(k)} w_{ij,l} - \theta_j\Big)$$

则隐层规模为 m 的 PNN 的隐层输出矩阵为

$$\boldsymbol{H}_m = [\boldsymbol{h}_1, \boldsymbol{h}_2, \cdots, \boldsymbol{h}_m]^{\mathrm{T}} \tag{9}$$

新增节点与现有节点相关性判别可描述为定理 1：

定理 1　记 PNN 隐层新增临时节点序号为 $m+1$，该节点对训练样本集的输出向量为 $\boldsymbol{h}_{m+1}$，网络中各节点输出的 Gram-Schmidt 正交向量为 $\{\boldsymbol{s}_1, \boldsymbol{s}_2, \cdots, \boldsymbol{s}_{\mathrm{m}}, \boldsymbol{s}_{m+1}\}$，则当 $\|\boldsymbol{s}_{m+1}\|_2 = 0$ 时，新增临时节点 $m+1$ 与相邻隐层节点 m 线性无关.

证明　设 $\boldsymbol{h}_{m+1}$ 为隐层新增临时节点的输出向量，根据 Gram-Schmidt 正交化法，隐层现有节点与新增临时节点的输出正交向量分别为

$$\begin{aligned}
&\boldsymbol{s}_1 = \boldsymbol{h}_1\\
&\boldsymbol{s}_2 = \boldsymbol{h}_2 - \frac{[\boldsymbol{s}_1, \boldsymbol{h}_2]}{[\boldsymbol{s}_1, \boldsymbol{s}_1]}\boldsymbol{s}_1\\
&\vdots\\
&\boldsymbol{s}_m = \boldsymbol{h}_m - \frac{[\boldsymbol{s}_1, \boldsymbol{h}_m]}{[\boldsymbol{s}_1, \boldsymbol{s}_1]}\boldsymbol{s}_1 - \frac{[\boldsymbol{s}_2, \boldsymbol{h}_m]}{[\boldsymbol{s}_2, \boldsymbol{s}_2]}\boldsymbol{s}_2 - \cdots - \\
&\qquad \frac{[\boldsymbol{s}_{m-1}, \boldsymbol{h}_m]}{[\boldsymbol{s}_{m-1}, \boldsymbol{s}_{m-1}]}\boldsymbol{s}_{m-1}\\
&\boldsymbol{s}_{m+1} = \boldsymbol{h}_{m+1} - \frac{[\boldsymbol{s}_1, \boldsymbol{h}_{m+1}]}{[\boldsymbol{s}_1, \boldsymbol{s}_1]}\boldsymbol{s}_1 - \frac{[\boldsymbol{s}_2, \boldsymbol{h}_{m+1}]}{[\boldsymbol{s}_2, \boldsymbol{s}_2]}\boldsymbol{s}_2 - \cdots - \\
&\qquad \frac{[\boldsymbol{s}_m, \boldsymbol{h}_{m+1}]}{[\boldsymbol{s}_m, \boldsymbol{s}_m]}\boldsymbol{s}_m
\end{aligned} \tag{10}$$

为表述方便，记 $\Lambda(\cdot)$ 表示新增节点与相邻节点的输出向量线性相关，$\Psi(\cdot)$ 为线性无关.

1）根据线性理论，当 $\Psi(\boldsymbol{h}_m, \boldsymbol{h}_{m+1})$ 时，$\Psi(\boldsymbol{s}_m, \boldsymbol{s}_{m+1})$ 成立，即 $\mathrm{span}(\boldsymbol{h}_1, \boldsymbol{h}_2, \cdots, \boldsymbol{h}_n) = \mathrm{span}(\boldsymbol{s}_1, \boldsymbol{s}_2, \cdots, \boldsymbol{s}_n)$，其中 $\mathrm{span}(\boldsymbol{s}_1, \cdots, \boldsymbol{s}_n)$ 表示由 $(\boldsymbol{s}_1, \cdots, \boldsymbol{s}_n)$ 张成的向量空间.

2）当 $\Lambda(\boldsymbol{h}_m, \boldsymbol{h}_{m+1})$ 时，记 $\boldsymbol{h}_{m+1} = \alpha\boldsymbol{h}_m$，则有：

$$\begin{aligned}
\boldsymbol{s}_{m+1} &= \boldsymbol{h}_{m+1} - \sum_{i=1}^{m}\frac{[\boldsymbol{s}_i, \boldsymbol{h}_{m+1}]}{[\boldsymbol{s}_i, \boldsymbol{s}_i]}\boldsymbol{s}_i\\
&= \alpha\boldsymbol{h}_m - \sum_{i=1}^{m}\frac{[\boldsymbol{s}_i, \alpha\boldsymbol{h}_m]}{[\boldsymbol{s}_i, \boldsymbol{s}_i]}\boldsymbol{s}_i\\
&= \alpha\boldsymbol{h}_m - \alpha\sum_{i=1}^{m-1}\frac{[\boldsymbol{s}_i, \boldsymbol{h}_m]}{[\boldsymbol{s}_i, \boldsymbol{s}_i]}\boldsymbol{s}_i - \alpha\frac{[\boldsymbol{s}_m, \boldsymbol{h}_m]}{[\boldsymbol{s}_m, \boldsymbol{s}_m]}\boldsymbol{s}_m\\
&= \alpha\Big(\boldsymbol{s}_m - \frac{[\boldsymbol{s}_m, \boldsymbol{h}_m]}{[\boldsymbol{s}_m, \boldsymbol{s}_m]}\boldsymbol{s}_m\Big)\\
&= [0, 0, \cdots, 0]^{\mathrm{T}}
\end{aligned}$$

因此 $\|\boldsymbol{s}_{m+1}\|_2 = 0$. 证毕.

根据定理 1，通过计算新增临时节点的 Gram-Schmidt 正交化向量可快速分析出新增临时节点 $m+1$ 与相邻隐层节点 m 间的相关性. IE-PNN 通过逐次添加新节点，判断临时节点的相关性. 若线性相关，则丢弃该节点并重新生成临时节点，重复该过程直到达到最小输出误差.

2.3　新增节点的输出权值计算

根据极限学习理论，IE-PNN 固定隐层现有节点的权函数正交基展开系数和输出权值，对新增节点的权函数 $w_{i,m+1}(t)$ 展开系数 $w_{i,m+1}^{(1)}, w_{i,m+1}^{(2)}, \cdots, w_{i,m+1}^{(L)}$ 实施随机赋值，则新增节点的输出权值为

$$\boldsymbol{v}_{m+1} = \frac{[\boldsymbol{e}_m, \boldsymbol{h}_{m+1}]}{\|\boldsymbol{h}_{m+1}\|_2} \tag{11}$$

其中，$\boldsymbol{e}_m$ 为有 m 个隐层节点的 IE-PNN 输出误差：

$$\begin{aligned}
\boldsymbol{e}_m &= \boldsymbol{e}_{m-1} - \boldsymbol{v}_m\boldsymbol{h}_m\\
&= \boldsymbol{e}_{m-1} - \boldsymbol{v}_m[f(a_{i,l}^{(1)}, w_{im,l}, \theta_m), \cdots,\\
&\qquad f(a_{i,l}^{(K)}, w_{im,l}, \theta_m)]
\end{aligned} \tag{12}$$

根据新增节点输入权值参数优化、节点相关性判别和输出权值计算三个方面，实现 IE-PNN 的隐层结构自增长机制. 其收敛性可以表述为定理 2：

定理 2　给定一组独立同分布的学习样本集 $\{X_k(t), d_k\}_{k=1}^{K}$，则对于给定任意小的训练误差精度 ε，如果过程神经网络的参数由 IE-PNN 获得，则必存在一个隐层过程神经元节点数为 m，使得 $\|E_m(\boldsymbol{H})\| = \|\boldsymbol{H}_{K\times m}\boldsymbol{V}_{m\times q} - \boldsymbol{T}_{K\times q}\| \leqslant \varepsilon$ 成立.

证明　设过程神经网络的最大隐层节点数为 $\tilde{N}$，用于优化新增节点输入参数的 QFA 的最大优化次数为 G，当前优化次数为 g，当前网络隐层节点数为 λ.

1）当 $\lambda = 1$，$g = 0$ 时：

$$\|E_\lambda^g(\boldsymbol{H})\| = \|\boldsymbol{H}_{K\times\lambda}\boldsymbol{V}_{\lambda\times q} - \boldsymbol{T}_{K\times q}\| > \varepsilon$$

2）当 $\lambda = 2$ 时，根据 QFA 全局收敛性可得

$$\|E_2^G(\boldsymbol{H})\| \leqslant \cdots \leqslant \|E_2^2(\boldsymbol{H})\| \leqslant \|E_2^1(\boldsymbol{H})\| \leqslant \|E_1^0(\boldsymbol{H})\|$$

3）同理，当 $\lambda = m$ 时：

$$\begin{aligned}
\|E_m^G(\boldsymbol{H})\| &\leqslant \cdots \leqslant \|E_m^1(\boldsymbol{H})\| \leqslant \cdots \leqslant \|E_2^G(\boldsymbol{H})\|\\
&\leqslant \cdots \leqslant \|E_1^0(\boldsymbol{H})\|
\end{aligned}$$

4）根据单调递减特征，存在 m 使 $|E_m^g(\boldsymbol{H})\| \leqslant \varepsilon$ 成立.

以上过程说明，存在一个 m 且 $m < \tilde{N}$，网络收敛成立. 证毕.

3　Henon 时间序列预测

为验证所提方法有效性，与其它 PNN 模型对比，包括正交基展开 PNN(OBPNN)[2]、粒子群优化结构的 PNN(PSO-PNN)[7]、极限学习 PNN(FE-PNN)[17].

3.1　仿真实验

Henon 时间序列预测属于动力系统研究的反问题，常用来检验非线性系统模型的性能，其动力学方程为

$$\begin{cases} x(t+1) = 1 + y(t) - ax^2(t)\\ y(t+1) = bx(t) \end{cases}$$

其中，当 $a = 1.4$，$b = 0.3$ 时，Henon 系统进入混沌状态，此时仿真预测是动力学系统研究的经典问题.

1）样本构造和参数：系统初始条件为 $x(0) = 0.5$，$y(0) = 0.5$，取 $t \in [0, 149]$ 的 150 个数据，记为 $\{x(t)\}_{t=0}^{149}$. 按照数据的先后关系，连续 6 个构成一个输入向量，对第 7 个预测. 以此类推共得 144 组样本，训练样本选择其中

的 84 组，剩余 60 组为测试. 4 种过程神经网络模型的正交基函数均为 Legendre 函数，满足拟合精度时的函数项 $L=8$，训练误差精度 $\varepsilon=0.001$. 此外为方便对比，参照文[2]，OBPNN 的学习速度参数设置为 $\alpha=\beta=\gamma=0.5$，惯性系数 $\eta=0.1$.

此外，对用于 IE-PNN 输入权函数基展开系数优化的量子衍生萤火虫算法 QFA 中的光吸收系数 γ 取值进行讨论，根据文[19]，$\gamma\in[0.1, 100]$. 本文从实验分析角度讨论 γ 对 IE-PNN 训练的影响. 图 2 为 γ 不同取值时，量子编码萤火虫在单位圆的距离 r_{ij} 与相对吸引力 β_{ij} 的关系：

（1）当 $\gamma=0.5$ 时，萤火虫间的相对吸引力 β_{ij} 值随距离增加下降较为缓慢，最远处 $r_{ij}=\pi$ 时仍然大于 0.7，表明距离较远的萤火虫间也会产生较大的吸引力，有利于种群进化.

（2）当 $\gamma=10$ 时，β_{ij} 值随距离增加急剧减小，萤火虫对 $r_{ij}>0.3\pi$ 的其它萤火虫的吸引力趋于 0，此时进化过程较慢.

因此从吸引力角度看，选取较小的 γ 更有利于种群进化. 此外图 3 为 γ 与 IE-PNN 训练收敛时间 time 间的关系，表明当 γ 值过小时，萤火虫个体吸引力过大，种群易陷入局部极值，影响 IE-PNN 的学习速度. 综上所述，选取 $\gamma=0.5$ 作为 QFA 的光吸收系数.

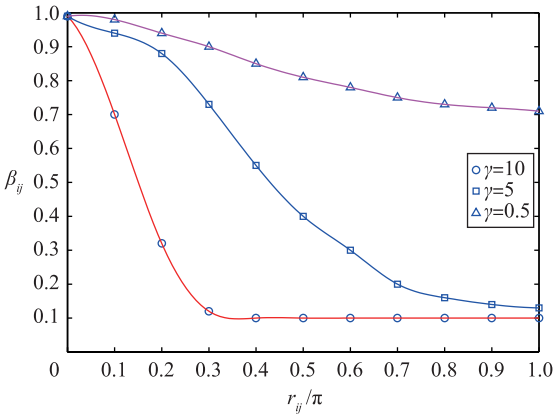

图 2 γ 取值对 $r_{ij}-\beta_{ij}$ 曲线的影响

Fig.2 Influence of γ value on the curve $r_{ij}-\beta_{ij}$

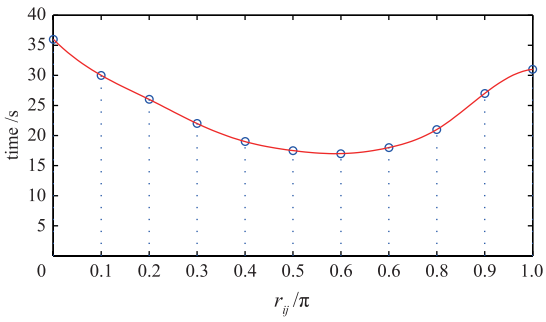

图 3 γ 取值对训练的影响

Fig.3 Influence of γ on the training process

2）仿真实验对比：分别对 IE-PNN 的学习和预测能力进行验证，其中学习能力指标包括训练耗时 time、训练收敛率 TCR、最大训练时间 Max_T 和最小训练时间 Min_T，其中 time 越小、TCR 越高，训练算法的学习能力越好. 预测能力指标包括预测的确定系数 E、平均绝对误差 MAE、平均相对误差 MRE、均方根误差 RMSE. 其中 E 越接近 1、RMSE 接近 0，预测效果越好. 针对 4 种模型和 84 组训练样本，各进行 20 次仿真实验，训练情况对比如表 1 所示. 相对于 OBPNN 和 PSO-PNN，IE-PNN 的训练时间有明显提高. 相对于 FE-PNN，虽然训练时间有一定增加，但训练收敛率提高近 35%.

表 1 4 种过程神经网络的训练对比

Tab.1 Training comparison among the four PNN models

模型	time /s	TCR	Min_T	Max_T
OBPNN	81.50	0.65	77.53	85.31
PSO-PNN	164.73	0.85	149.85	181.27
FE-PNN	6.22	0.55	5.97	6.65
IE-PNN	17.07	0.90	15.62	21.25

4 种模型在每次训练结束并收敛的情况下，立即对测试样本集进行预测，并计算各个预测指标，预测情况对比如表 2 所示. 从实验结果中可明显看出 IE-PNN 预测能力明显优于 OBPNN 和 FE-PNN，比 PSO-PNN 也有提高.

表 2 4 种过程神经网络的预测对比

Tab.2 Prediction comparison among the four PNN models

模型	E	MAE	MRE	RMSE
OBPNN	0.883 7	0.069 2	0.311 9	0.068 1
PSO-PNN	0.940 6	0.027 3	0.107 3	0.037 4
FE-PNN	0.913 1	0.035 5	0.184 0	0.040 6
IE-PNN	0.978 2	0.018 5	0.054 9	0.025 1

3.2 实验结果分析

1）训练时间从快到慢依次为：FE-PNN、IE-PNN、OBPNN 和 PSO-PNN. 原因分析：IE-PNN 和 FE-PNN 只进行向量矩阵计算，其中 FE-PNN 通过 Moore-Penrose 广义逆一次完成隐层所有节点的输出权值计算，而 IE-PNN 对每个新增节点实施相关性判别和输出误差计算，因此训练时间长于 FE-PNN，但 FE-PNN 隐层输出矩阵易出现病态特征，收敛性较差. 而 PSO-PNN 采用种群进化方式，对模型所有参数进行粒子编码，每次进化需要计算多组可行解的误差精度，相当于多个模型同时训练，所以 PSO-PNN 的时间最长. 对于输入输出结构为 $n-m-1$ 的模型结构，PSO-PNN 种群为 A，最大迭代次数为 B，IE-PNN 逐次新增节点时相关性节点数为 C，则 PSO-PNN 的时间复杂度为 $O(AB(nmL+m))$，IE-PNN 的时间复杂度仅为 $O((m+C)L)$.

2）预测能力从高到低依次为：IE-PNN、PSO-PNN、FE-PNN 和 OBPNN. 原因：IE-PNN 通过 Gram-Schmidt 正交化完成新增节点相关性判别，使得网络在节点不断增加过程中消除冗余节点，保持网络结构紧凑且每个节点的输出权值均为最小二乘解，因此模型的逼近效率最高. PSO-PNN 利用进化算法完成模型参数和结构的优化，但结构中仍存在冗余节点且训练过程属于高维优化问题，易陷入局部最小. FE-PNN 则是利用较多的隐层节点来提高模型的泛化

能力，隐层节点的数量对于 FE-PNN 较为重要，试探法确定隐层规模较为耗时. 而 OBPNN 梯度下降训练时对初值敏感，影响模型逼近效率.

综上所述，IE-PNN 训练效率优于 OBPNN 和 PSO-PNN，预测能力也明显高于 OBPNN，相对于 PSO-PNN 也有提高，并较好地解决了 FE-PNN 训练收敛性低的不足.

4　岩性识别应用

岩性识别页岩油测井评价的前提和基础，是页岩油勘探开发的一项重要研究内容. 本文的岩性样本数据来源于罗家地区，测井资料包括自然伽马 GR、中子 CNL、声波时差 AC、密度 DEN、自然电位 SP、井径 CAL、深侧向电阻率 RLLD、浅侧向电阻率 RLLS，岩性类别包括泥岩、砂岩、灰岩、油页岩等.

1）数据归一化：不同测井曲线量纲不同，因此要对测井曲线做归一化处理，消除量纲不同带来的影响，具体为

$$X^* = \frac{X - X_{\min}}{X_{\max} - X_{\min}}$$

其中，X、$X_{\min}$、$X_{\max}$和 X^* 都为测井曲线离散数据，其中 X 为原始值，$X_{\min}$和 $X_{\max}$是最小值和最大值，X^* 为归一化后数值. 此外由于电阻率是具有非线性对数特征的测井曲线，归一化公式为

$$X^* = \frac{\lg X - \lg X_{\min}}{\lg X_{\max} - \lg X_{\min}}$$

2）PCA 特征提取：由于测井曲线是岩性的综合反映，选取测井曲线越多，越有利于提高岩性识别准确率. 但当特征输入较多时，神经网络模型不仅计算效率低、收敛速度慢，而且还会因输入维数过高引起网络结构复杂化，降低模型泛化能力. 因此本文首先使用主成分分析法 PCA 对所选测井曲线实施主成分分析，表 3 为 8 条测井曲线的主成分统计分析结果，其中前 3 个主成分的累积贡献率达到 94.982 9%，超过了 85%，因此最终选用前 3 个主成分的 PCA 统计分析得分指标来表示原数据集中的 8 条测井曲线.

表 3　测井曲线主成分统计分析

Tab.3　PCA statistical analysis for logging curves

成分	特征根	方差贡献率 /%	累积贡献率 /%
1	1.557 3	59.821 9	59.821 9
2	0.757 3	29.089 8	88.911 8
3	0.158 0	6.071 1	94.982 9
4	0.088 0	3.380 4	98.363 3
5	0.027 7	1.065 2	99.428 5
6	0.014 9	0.571 5	100.000 0

3）岩性识别结果：根据 A1 井、A2 井、A3 井的岩性剖面和测井资料，共选取 210 个岩性识别样本组成训练样本集，样本分布为泥岩 50 个、灰岩 60 个、油泥岩 60 个、油页岩 40 个，每个样本为 4 个长度为 10 的离散序列. 当网络训练结束后，以 A4 井 3 080 m ~ 3 140 m 的井段为例，验证模型的岩性识别能力. 其中参与对比的 4 个神经网络模型，岩性识别指标对比如表 4 所示. 其中本文所提的 IE-PNN 的岩性识别率高于其它 3 种 PNN 模型，相对于文[17]的 FE-PNN，也提高了 5 个百分点.

表 4　岩性识别的结果对比

Tab.4　Result comparison about lithology recognition

模型	识别率 /%				平均识别率 /%
	油页岩	油泥岩	泥岩	灰岩	
OBPNN	66.67	82.35	60.00	83.33	75.00
PSO-PNN	76.47	88.24	75.00	100.00	83.33
FE-PNN	77.78	88.24	91.67	100.00	86.11
IE-PNN	88.89	94.12	91.67	100.00	91.67

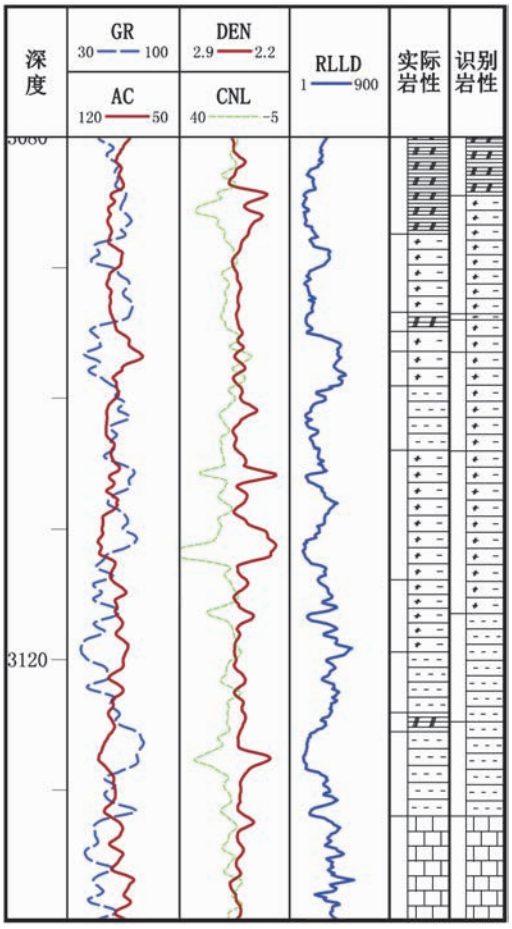

图 4　岩性识别的垂向分布

Fig.4　Vertical distribution of the lithology identification

图 4 是 IE-PNN 模型对罗家地区的 A4 井 3 080 m ~ 3 140 m井段的岩性识别垂向分布情况，对比的真实岩性来自于实际的取芯资料. 从图 4 中的识别岩性和实际岩心的对比可以明显看出 IE-PNN 获得了较好的岩性分类结果，其中对于灰岩和油泥岩的分类效果最为理想，油页岩的分类精度也达到 88% 左右. 该结果表明 IE-PNN 可识别油藏地质中的复杂岩性，提高分类精度. 此外，当未使用 PCA 时，IE-PNN 在该井段的岩性识别率仅为 84% 左右，相对目前的方法降低了 7 个百分点. 该结果表明本文对测井曲线使用 PCA 特征提取和降维的有效性，不仅消除测

井曲线数据间的自相关性，并实现数据压缩和降维，加快网络收敛速度，提高了网络识别精度.

5 结论

过程神经网络模型由于自身的复杂性，隐层节点规模的选取目前主要采用试探法，从而增加了模型的复杂性，同时也降低了 PNN 自身的逼近效率. 本文主要针对过程神经网络的隐层结构自增长机制和模型训练算法两个方面，开展对增量极限学习过程神经网络的研究. 通过对新增节点的输入权值优化、相关性分析、输出权值计算和收敛性分析四个方面，建立隐层结构自增长机制，从而消除模型中的冗余节点，提高结构的紧凑性，最后通过仿真实验和实际工程应用验证了模型的有效性，从而为过程神经元网络的结构自适应研究提供了一种新的思路.

参考文献

[1] He X G, Liang J Z. Some theoretical issues on procedure neural networks[J]. Engineering Science, 2000, 2(12): 40 – 44.

[2] 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[J]. 计算机学报，2004，27(5)：645 – 650.
Xu S H, He X G. Learning algorithms of process neural networks based on orthogonal function basis expansion[J]. Chinese Journal of Computers, 2004, 27(5): 645 – 650.

[3] Luan S G, Zhong S S, Li Y. Hybrid recurrent process neural network for aero engine condition monitoring[J]. Neural Network World Journal, 2008, 2(18): 133 – 145.

[4] Cheng M, Ye S. Fuzzy process neural network based on orthogonal basis function[J]. Information Technology Journal, 2011, 10(10): 1999 – 2003.

[5] Li P C, Shi G Y. Numerical integration-based discrete process neural networks algorithm and applications[J]. System Engineering - Theory & Practice, 2013, 33(12): 3216 – 3222.

[6] 李盼池，王海英，戴庆，等. 量子过程神经网络模型算法及应用[J]. 物理学报，2012，61(16)：160303 –1 –160303 –9.
Li P C, Wang H Y, Dai Q, et al. Quantum process neural networks model algorithm and applications[J]. Acta Physica Sinica, 2012, 61(16): 160303 –1 –160303 –9.

[7] 刘坤，谭营，何新贵. 基于粒子群优化的过程神经网络学习算法[J]. 北京大学学报(自然科学版)，2011，47(2)：238 – 244.
Liu K, Tan Y, He X G. Particle swarm optimization based learning algorithm for process neural networks[J]. Acta Scientiarum Naturalium Universitatis Pekinensis, 2011, 47(2): 238 – 244.

[8] Huang G B, Zhu Q Y, Siew C K. Extreme learning machine: A new learning scheme of feedforward neural networks[C]//International Joint Conference on Neural Networks. Piscataway, NJ, USA: IEEE, 2004: 985 – 990.

[9] Huang G B. Extreme learning machine for regression and multiclass classification[J]. IEEE Transactions on Systems, 2012, 42(2): 513 – 516.

[10] Huang Z Y, Yu Y L, Gu J, et al. An efficient method for traffic sign recognition based on extreme learning machine[J]. IEEE Transactions on Cybernetics, 2017, 47(4): 920 – 933.

[11] Huang G, Song S, Gupta J N. Semi-supervised and unsupervised extreme learning machines[J]. IEEE Intelligent System, 2014, 44(12): 2405 – 2417.

[12] Kasun L L C, Zhou H M, Huang G B. Representational learning with extreme learning machine for big data[J]. IEEE Intelligent System, 2013, 28(6): 31 – 34.

[13] Huang G B, Zhou H, Ding X, et al. Extreme learning machine for regression and multiclass classification[J]. IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics, 2012, 42(2): 513 – 529.

[14] 马超，张英堂. 在线核极限学习机及其在时间序列预测中的应用[J]. 信息与控制，2014，43(5)：624 – 629.
Ma C, Zhang Y T. Online kernel extreme learning machine and its application to time series prediction[J]. Information and Control, 2014, 43(5): 624 – 629.

[15] Chen Y, Yao E, Basu A. A 128-channel extreme learning machine-based neural decoder for brain machine interfaces[J]. IEEE Transactions on Biomed Circuits Systems, 2015, 10(3): 679.

[16] Wang Y, Yu H, Huang G B, et al. An energy-efficient nonvolatile in-memory computing architecture for extreme learning machine by domain-wall nanowire devices[J]. IEEE Transactions on Nanotechnologies, 2015, 14(6): 998 – 1012.

[17] 刘志刚，李盼池，许少华. 基于 ELM-PNN 算法的第 24 周太阳黑子预测预报[J]. 控制与决策，2017，32(4)：642 – 646.
Liu Z G, Li P C, Xu S H. Prediction and forecast of sunspot numbers in 24 cycle based on ELM-PNN algorithms[J]. Control and Decision, 2017, 32(4): 642 – 646.

[18] 刘志刚，许少华，李盼池. 基于 QPSO 和极限学习的离散过程神经网络及学习算法[J]. 控制与决策，2016，31(12)：2241 – 2247.
Liu Z G, Xu S H, Li P C. Discrete process neural networks and learning algorithms based on QPSO and extreme learning machine[J]. Control and Decision, 2016, 31(12): 2241 – 2247.

[19] Yang X S. Nature-inspired metaheuristic algorithms[M]. Frome, UK: Luniver Press, 2008: 81 – 89.

[20] Yang X S. Firefly algorithm, stochastic test functions and design optimization[J]. International Journal of Bio-Inspired Computation, 2010, 2(2): 78 – 84.

作者简介

杜　娟(1980 –)，女，硕士，副教授. 研究领域为过程神经网络，极限学习机.

刘志刚(1979 –)，男，博士，副教授. 研究领域为极限学习机，进化算法.

许少华(1962 –)，男，博士生导师，教授. 研究领域为过程神经网络，智能计算.