

# 基于 ELM 神经网络的 FAST 节点位移预测研究

沙 毅<sup>1</sup>, 陈 曦<sup>1</sup>, 张立立<sup>1</sup>, 朱丽春<sup>2</sup>  
(1. 东北大学 计算机科学与工程学院, 辽宁 沈阳 110169; 2. 中国科学院 国家天文台, 北京 100012)

**摘 要:** 针对 ELM 神经网络隐含层节点数目需要人工设定, 容易出现过拟合现象从而导致网络的泛化能力降低的问题, 引出了基于误差最小化的 ELM 神经网络的改进方法 EM\_ELM 算法, 并在理论上论证了 EM\_ELM 算法对于提高 ELM 神经网络预测精度和泛化能力的可行性. 随后将 EM\_ELM 算法应用到 FAST 节点位移的预测模型中, 并且进行了仿真验证. 仿真结果表明虽然 EM\_ELM 神经网络在训练时间上有了一定的损失, 但是仍能满足实时性的要求, 而且它的预测精度和泛化能力都得到提升, 证明了改进算法的有效性与可行性, 进一步说明了 EM\_ELM 神经网络更适合应用于 FAST 节点位移预测.

**关 键 词:** FAST 节点; ELM; 神经网络; 位移预测; 可行性

中图分类号: TP 393.17      文献标志码: A      文章编号: 1005-3026(2017)05-0630-04

## Research on FAST Node Displacement Prediction Based on ELM Neural Network

SHA Yi<sup>1</sup>, CHEN Xi<sup>1</sup>, ZHANG Li-li<sup>1</sup>, ZHU Li-chun<sup>2</sup>  
(1. School of Computer Science & Engineering, Northeastern University, Shenyang 110169, China; 2. National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012, China. Corresponding author: SHA Yi, E-mail: shayi@ise.neu.edu.cn)

**Abstract:** Due to the problems that the numbers of nodes in hidden layers of ELM neural network are in need of manual setting, and the over-fitting phenomenon is easy to appear, resulting in a reduction in the network generalization, an EM\_ELM algorithm was proposed to improve ELM neural network based on error minimization. The feasibility was proved in theory which could improve the prediction accuracy and generalization of ELM neural network. Meanwhile, the algorithm was also applied into the model of FAST node displacement prediction and conducted simulation finally. The results show that although EM\_ELM neural network is not sufficient in training time to a certain degree, it is still proper in real-time requirement. Besides, its prediction accuracy and generalization capabilities are enhanced, which is just a proof in the effectiveness and feasibility of the improved algorithm, thereby further illustrating that the EM\_ELM neural network is more suitable for FAST node displacement prediction.

**Key words:** FAST node; ELM; neural network; displacement prediction; feasibility

FAST 工程<sup>[1]</sup>实施至今进展顺利, 而 FAST 主动反射面整网变形控制策略与自适应建模<sup>[2]</sup>研究作为 FAST 项目的基础工作和核心难点之一也取得了一定的进展. 主动反射面是 500 m 口径球面射电望远镜 FAST 创新工程的核心, 整网变形<sup>[3]</sup>过程中索网节点位置的实时、动态、精准控制<sup>[4]</sup>是保证望远镜观测效果的重要前提. 对反射面变形过程中 FAST 节点位移预测模型的研究和探索对于工业控制领域和 FAST 项目的设计<sup>[5]</sup>与实施具有很好的理论意义和实践价值.

ELM 神经网络又名极限学习机 (extreme learning machine, ELM) 是由 Huang 等<sup>[6]</sup>在 2004 年提出的一种单隐层前馈神经网络 (single-hidden layer feedforward neural networks, SLFNs) 的快速

学习算法<sup>[7]</sup>. 尽管单隐层前馈神经网络的学习能力很强,但是它也存在一定的缺陷,就是学习速度比较慢,无法满足一些工程研究中的要求. ELM 算法的结构与单隐层前馈神经网络有些相似,但 ELM 应用更为便捷,它在训练参数的选择方面比较灵活,而且预获得完整的网络训练模型也只需通过最小二乘的方法求出输出权值即可. Huang 等证明了神经网络隐含层节点参数的随机选取<sup>[8]</sup>,在不影响网络收敛能力的前提下,可以大大缩短网络训练的时间. ELM 神经网络的训练速度与传统前馈神经网络,如 BP 神经网络和 SVM 等相比,提高了数百倍到数千倍,引起了广大学者的极大兴趣.

# 1 极限学习机(ELM)算法

单隐层前馈型网络<sup>[9]</sup>包含三层,分别是输入层、隐含层和输出层. ELM 算法包含以下三个步骤.

ELM 算法:给定训练集  $D = \{(\mathbf{x}_i, \mathbf{t}_i) \mid \mathbf{x}_i \in \mathbf{R}^n, \mathbf{t}_i \in \mathbf{R}^m, i = 1, \cdots, N\}$ , 激活函数  $g(\mathbf{x})$ ,  $L$  为隐含节点个数.

步骤 1 随机生成隐单元的输出权值和偏置  $\mathbf{a}_i, \mathbf{b}_i, i = 1, \cdots, L$ ;

步骤 2 计算隐单元的输出矩阵  $\mathbf{H}$ ;

步骤 3 计算出权值  $\hat{\beta}; \hat{\beta} = \mathbf{H}^+ \mathbf{T}$ ;

其中,  $\mathbf{H}^+$  是隐含层输出矩阵  $\mathbf{H}$  的摩尔 - 彭罗斯广义逆,在 ELM 中,可以用奇异值的分解法来计算摩尔 - 彭罗斯广义逆.

$$\hat{\beta} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{T}. \tag{1}$$

其中,  $\mathbf{H}^+ = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T$ , 因为  $\mathbf{H}^+ \mathbf{H} = \mathbf{I}$ , 所以  $\mathbf{H}^+$  是  $\mathbf{H}$  的左伪逆. 因此,线性系统的一般最小均方估计为

$$\hat{\beta} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{T}. \tag{2}$$

ELM 神经网络与传统的前馈神经网络相比,其最大优点是其训练速度的优势非常明显. 因为 ELM 神经网络的初始隐含层的参数是随机生成的,在神经网络训练过程中参数不改变,通过持续的迭代,即可得出输出权值. 但是 ELM 神经网络也有一些不足之处. 不足之处体现在人工设置隐含层节点数目,通过大量实验发现,这种操作会在一定程度上对网络的性能造成不好的影响.

但是在实际应用中,往往都是通过大量实验数据确定隐含层节点的数目,如果隐含层节点数目选择过多,在神经网络训练的过程中会出现过拟合现象,使训练精度和测试精度不一致,最终导

致网络的鲁棒性和泛化性降低,不能满足在实际应用中的要求. 为了解决隐含层节点上述问题,进一步提高网络的性能,需要一种具有削减隐含层节点功能的改进 ELM 神经网络,来提高网络性能的参数.

# 2 ELM 的改进算法

## 2.1 基于误差最小化的 ELM 神经网络

ELM 神经网络的网络结构是由用户设定,若设置的网络结构复杂,会导致网络学习时间增加甚至出现过拟合等现象,会大大降低神经网络预测的准确性,当简单设置网络结构时,又会导致网络的训练误差较大. 现有的 ELM 神经网络的研究中,在大多数情况下,网络隐含节点数的选取也只能依赖于用户多次尝试后取经验值. 针对这些问题,很多研究人员尝试给出优化算法,其中增长型的 ELM 神经网络是其中行之有效的算法. 其算法不需要构建网络结构,首先将隐含节点的个数从 0 开始增加,随着隐含层节点数量的增多,可以减小网络训练误差,直至训练误差达到用户期望值为止.

本文尝试在增长型 ELM 神经网络的基础上,引入基于误差最小化的 ELM 神经网络(error minimized ELM, EM\_ELM). EM\_ELM 神经网络算法的基本思想:①首先设置初始网络:隐含层节点数为 0;或者由用户自己定义初始网络结构. ②其次利用训练数据对神经网络进行训练,得到初始输出权重值和训练误差,通过不断地进行迭代,来增加隐含层节点数目. ③当到达一定的节点数目之后,将其增加到网络中,再更新整个网络的输出权重和训练误差. ④通过不停地迭代增加节点数目,直到节点数达到设置值,或者训练误差达到设计要求为止.

## 2.2 EM\_ELM 算法步骤

EM\_ELM 算法中,初始网络结构由用户设定,隐含层节点数目通过线性方式增加, EM\_ELM 算法的具体流程如下:给定一个具有  $N$  个不同训练样本  $\{(\mathbf{x}_i, \mathbf{t}_i)\}_{i=1}^N$  的训练集,隐含层映射函数为  $G(\mathbf{x})$ , 隐含层的最大节点数为  $L_{\max}$ , 初始隐含层节点数为  $L_0$ , 用户期望训练误差为  $U$ .

1) 在输入权值和阈值的取值范围内为  $L_0$  个隐含层节点随机产生相应的输入权重和阈值.

2) 计算得到初始网络的输出矩阵  $\mathbf{H}_0$ :

$$\begin{aligned} & \mathbf{H}_0(\mathbf{a}_1, \cdots, \mathbf{a}_L, \mathbf{b}_1, \cdots, \mathbf{b}_L, \mathbf{x}_1, \cdots, \mathbf{x}_N) = \\ & \begin{bmatrix} G(\mathbf{a}_1, \mathbf{b}_1, \mathbf{x}_1) & \cdots & G(\mathbf{a}_L, \mathbf{b}_L, \mathbf{x}_1) \\ \vdots & & \vdots \\ G(\mathbf{a}_1, \mathbf{b}_1, \mathbf{x}_N) & \cdots & G(\mathbf{a}_L, \mathbf{b}_L, \mathbf{x}_N) \end{bmatrix}_{N \times L_0}. \end{aligned} \quad (3)$$

3) 计算初始网络的训练误差  $\mathbf{E}_0$ :

$$\mathbf{E}_0 = \|\mathbf{H}_0 \mathbf{H}_0^T \mathbf{T} - \mathbf{T}\|. \quad (4)$$

4) 记迭代次数  $k=0$ .

5) 判断网络的剩余误差是否小于用户期望训练误差  $U$  或者网络的隐含层节点数是否达到预设的最大值,如果是则停止迭代,保存已经训练好的网络,否则继续执行 6).

6) 令  $k=k+1$ .

7) 根据式(5)计算增加更新后的隐含层节点数目:

$$L_k = L_{k-1} + \delta L_{k-1}. \quad (5)$$

8) 为新增加的  $\delta L_{k-1}$  个隐含层节点获得相应的输入权值向量  $\mathbf{a}_i$  和偏置  $\mathbf{b}_i$ ,由此得到新增隐含层节点的输出矩阵为  $\delta \mathbf{H}_{k-1}$ ,那么整个网络的隐含层输出矩阵  $\mathbf{H}_k$  可以表示为  $\mathbf{H}_k = [\mathbf{H}_{k-1}, \delta \mathbf{H}_{k-1}]$ .

$$\begin{aligned} & \delta \mathbf{H}_{k-1} = \\ & \begin{bmatrix} G(\mathbf{a}_{L_{k-1}+1}, \mathbf{b}_{L_{k-1}+1}, \mathbf{x}_1) & \cdots & G(\mathbf{a}_{L_k}, \mathbf{b}_{L_k}, \mathbf{x}_1) \\ \vdots & & \vdots \\ G(\mathbf{a}_{L_{k-1}+1}, \mathbf{b}_{L_{k-1}+1}, \mathbf{x}_N) & \cdots & G(\mathbf{a}_{L_k}, \mathbf{b}_{L_k}, \mathbf{x}_N) \end{bmatrix}_{N \times \delta L_{k-1}}. \end{aligned} \quad (6)$$

9) 计算整个网络的输出权值,

$$\mathbf{D}_k = ((\mathbf{I} - \mathbf{H}_k \mathbf{H}_k^T) \delta \mathbf{H}_k)^{\dagger}, \quad (7)$$

$$\mathbf{U}_k = \mathbf{H}_k^{\dagger} (\mathbf{I} - \delta \mathbf{H}_k^T \mathbf{D}_k), \quad (8)$$

$$\beta_k = \mathbf{H}_k^{\dagger} \mathbf{T} = \begin{bmatrix} \mathbf{U}_k \\ \mathbf{D}_k \end{bmatrix} \mathbf{T}. \quad (9)$$

10) 计算新增节点之后的训练误差  $\mathbf{E}_k$ ,然后跳转到 5) 继续执行,

$$\mathbf{E}_k = \|\mathbf{H}_k \beta_k - \mathbf{T}\|. \quad (10)$$

### 3 仿真结果及分析

经过多次试验,最终设定隐含层映射函数  $G(x)$  为高斯函数,初始隐含层节点数  $L_0$  为 10,隐含层的最大节点数  $L_{\max}$  为 100,用户期望训练误差  $U$  为 0.001.使用 200 组数据对 EM\_ELM 神经网络进行训练,训练耗时为 3.107 487 s.

将 EM\_ELM,ELM 和 RBF 神经网络<sup>[10]</sup>对 FAST 节点位移预测结果的各项性能参数列于表

1 中,明显可以看出虽然 EM\_ELM 神经网络的训练时间较 ELM 有所损失,但是仍然远优于 RBF 神经网络,而在预测结果的准确性和稳定性上 EM\_ELM 神经网络较 ELM 神经网络在平均绝对误差、平均相对误差和均方误差 3 个方面都有所提升,更适合应用于 FAST 中进行节点位移的预测.

表 1 EM\_ELM 与 ELM 及 RBF 神经网络的预测性能对比  
Table 1 Prediction performance comparison of EM\_ELM,ELM and RBF neural network

算法	训练时间 s	平均绝对 误差 mm	平均相 对误差 %	均方误差 mm
RBF	26.987 194	7.112 7	2.97	8.092 0
ELM	0.896 670	8.322 4	3.87	10.582 3
EM_ELM	3.107 487	4.479 8	2.38	5.085 8

预测结果的拟合曲线如图 1 和图 2 所示,预测结果的绝对误差如图 3 所示.

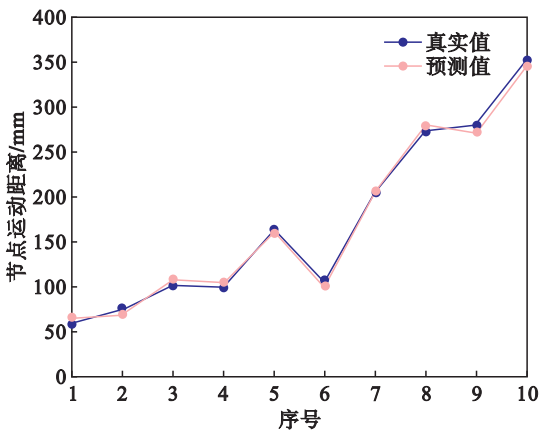


图 1 基于 EM\_ELM 神经网络的第一组节点位移预测结果  
Fig. 1 Prediction results of the first group nodes based on EM\_ELM neural network

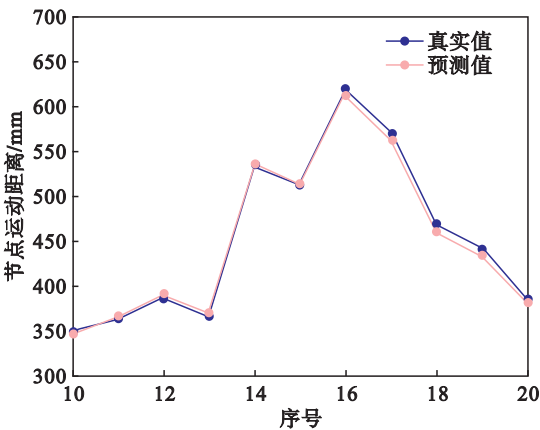


图 2 基于 EM\_ELM 神经网络的第二组节点位移预测结果  
Fig. 2 Prediction results of the second group nodes based on EM\_ELM neural network

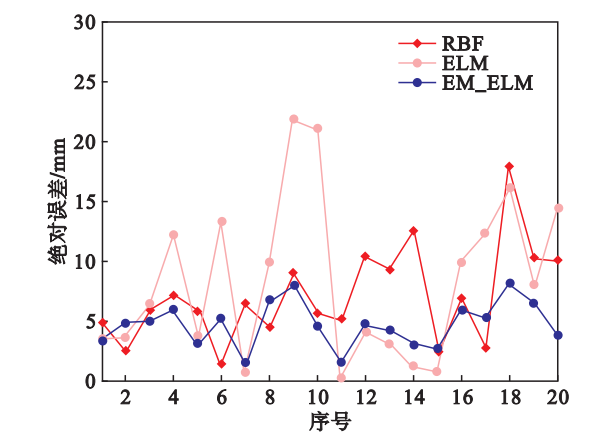


图 3 EM\_ELM 神经网络预测结果的绝对误差  
Fig. 3 Absolute error of the prediction results based on EM\_ELM neural network

从图 1 和图 2 中可以看出,EM\_ELM 算法的预测结果与真实位移相差不大,达到了预期的预测精度.从图 3 的误差对比可以得到,EM\_ELM 算法相比 RBF 和 ELM 算法,绝对误差波动范围较小,位移预测结果相对稳定,有利于算法在 FAST 预测模型中的应用.

# 4 结 论

本文介绍了 ELM 神经网络的理论基础,在此基础上对 ELM 算法的优缺点进行了分析.针对 ELM 神经网络隐含层节点数目需要人工设定,容易出现过拟合现象从而导致网络的泛化能力降低的问题,引出了 ELM 神经网络的改进方法 EM\_ELM 算法,并详细阐述了算法的实现过程.随后将 EM\_ELM 算法应用到 FAST 节点位移的预测模型中,并且进行了仿真实验,最终证明了改进算法的有效性与可行性.

# 参考文献:

[ 1 ] Qian H L, Fan F, Shen S Z. The cable-net structure supporting the reflector of FAST [ J ]. *China Civil Engineering Journal*, 2005, 38( 12 ): 18 - 23.

[ 2 ] 朱丽春. 500 米口径球面射电望远镜( FAST )主动反射面整网变形控制[ J ]. 科研信息化技术与应用, 2012, 3( 4 ): 67 - 75.  
( Zhu Li-chun. Control of the main active reflector of FAST [ J ]. *E-Science Technology & Application*, 2012, 3( 4 ): 67 - 75. )

[ 3 ] Jiang P, Wang Q M, Zhao Q. Optimization and analysis on cable net structure supporting the reflector of large radio telescope FAST[ J ]. *Engineering Mechanics*, 2013, 30( 2 ): 400 - 405.

[ 4 ] 杜敬利, 保宏, 杨东武, 等. 索网主动反射面的形状精度调整研究[ J ]. 工程力学, 2012, 29( 3 ): 212 - 217.  
( Du Jing-li, Bao Hong, Yang Dong-wu, et al. Analysis on shape accuracy adjustment of an active cable-mesh reflector [ J ]. *Engineering Mechanics*, 2012, 29( 3 ): 212 - 217. )

[ 5 ] 范峰, 牛爽, 钱宏亮, 等. FAST 背架结构优化选型及单元足尺模型试验研究[ J ]. 建筑结构学报, 2010, 31( 12 ): 9 - 16.  
( Fan Feng, Niu Shuang, Qian Hong-liang, et al Structural optimization and full-scale-model test for FAST back-structure unit [ J ]. *Journal of Building Structures*, 2010, 31( 12 ): 9 - 16. )

[ 6 ] Huang G B, Zhu Q Y, Siew C K. Extreme learning machine: a new learning scheme of feedforward neural networks [ C ] // IEEE International Conference on Neural Networks. Budapest, 2004: 985 - 990.

[ 7 ] Wang Y, Tian G, Wen S. Prediction model of television program rating based on BP neural network [ J ]. *Video Engineering*, 2014, 38( 6 ): 94 - 96.

[ 8 ] Huang G B, Zhu Q Y, Siew C K. Extreme learning machine: theory and applications [ J ]. *Neurocomputing*, 2006, 70( 1/2/ 3 ): 489 - 501.

[ 9 ] Soria-Olivas E, Gomez-Sanchis J, Martin J D, et al. BELM: Bayesian extreme learning machine [ J ]. *IEEE Transactions on Neural Networks*, 2011, 22( 3 ): 505 - 509.

[ 10 ] Gan M, Peng H, Dong X P. A hybrid algorithm to optimize RBF network architecture and parameters for nonlinear time series prediction[ J ]. *Applied Mathematical Modelling*, 2012, 36( 7 ): 2911 - 2919.