

一种新的小波神经网络结构优化设计方法

李益国, 沈 炯, 吕震中
(东南大学 动力工程系, 江苏 南京 210096)

摘要: 针对基于框架的小波神经网络存在的网络冗余性较大的问题, 提出一种基于粗糙集理论的网络结构优化设计方法. 首先通过时频分析确定小波神经网络的初步构造; 在此基础上, 根据网络输出对隐层节点依赖度的大小去除冗余的隐层节点, 达到优化网络结构的目的. 仿真结果表明该方法是简单而有效的.

关键词: 粗糙集理论; 小波神经网络; 小波框架; 属性依赖度

中图分类号: TP18 **文献标识码:** A

New approach to structure optimization of wavelet neural network

LI Yi-guo, SHEN Jiong, LÜ Zhen-zhong

(Department of Power Engineering, Southeast University, Jiangsu Nanjing 210096, China)

Abstract: Based on the rough sets theory, an approach was presented to minimize the redundancy of structure existing in wavelet neural networks. The original structure of wavelet network is obtained through time-frequency analysis. Then the redundant nodes are eliminated in light of dependency between the output of the network and the nodes in the hidden layers to optimize the structure of wavelet network. Simulation results illustrate the proposed method is simple and effective.

Key words: rough sets; wavelet neural network; wavelet frame; dependency of attributes

1 引言 (Introduction)

小波神经网络是近年来在小波分析研究获得突破的基础上提出的一种前馈型网络, 最早是由 Qinghua Zhang 于 1992 年在文献[1]中提出的. 由于小波神经网络与常规的神经网络相比具有很多优点, 如其结构和参数的确定借助小波分析理论有一定的指导, 网络权值学习算法简单、收敛速度快等, 因此最近被广泛应用于函数学习^[2]、过程控制^[3]等领域, 并取得了良好效果.

已有的小波神经网络基本上可以分为 3 类, 它们分别建立在连续小波变换、小波框架和正交小波变换理论的基础上. 由小波分析理论可知这 3 类小波神经网络都具有一致逼近和 L^2 逼近能力. 对于连续小波变换神经网络, 由于尺度和平移参数均可调, 使其与输出为非线性关系, 易带来类似 BP 网络的局部极小值等问题; 对于正交小波变换神经网络, 尽管理论上研究较为方便, 但正交基函数构造复杂; 而基于框架的小波神经网络, 由于网络权系数与输出呈线性关系, 不存在局部极小值问题, 同时基函数选取灵活, 因此小波框架神经网络更为实用.

基于框架的小波神经网络存在的最大问题是: 由于框架可以是线性相关的, 因此使用在时频平面上被正规截取的小波框架中的小波基函数作为隐层节点的神经元函数, 常常导致生成的网络具有很大的冗余性, 因此进行小波框架神经网络的结构优化具有重要意义. 文献[4]基于自适应投影算法提出一种小波神经网络的结构优化方法, 但该方法相对较复杂. 文献[5]提出的后向删除法仅适用于网络权值相差较大的情况.

粗糙集理论是一种刻划不完整性和不确定性的数学工具, 善于从大量数据中发现隐含的知识, 揭示潜在的规律. 近来已经有文献基于粗糙集理论进行 BP 网络设计的研究^[6]. 本文基于粗糙集理论提出一种小波神经网络的结构优化设计方法, 仿真结果表明该方法是有效的.

2 基于框架的小波神经网络 (Wavelet neural networks based on frame)

2.1 小波分析基础 (Basis of wavelet analysis)

设 $\Psi(x) \in L^2(\mathbb{R})$ 为小波函数, 对其进行平移和展缩变换, 生成一族小波函数

$$\Psi_{mn}(x) = a^{n/2}\Psi(a^n x - mb). \quad (1)$$

其中 $a, b \in \mathbb{R}, a > 1, b > 0, m, n \in \mathbb{Z}$, 则存在常数 $a_0 > 1, b_0 > 0$, 使得当 $a \in (1, a_0], b \in (0, b_0]$ 时, $\{\Psi_{mn}\}$ 构成 $L^2(\mathbb{R})$ 中的框架. 则对于任意能量有限信号 $f(x)$, 有下列展开式

$$f(x) = \sum c_{m,n}\Psi_{m,n}. \quad (2)$$

2.2 基于框架的小波神经网络结构 (Structure of wavelet neural networks based on frame)

式(2)可以由一个3层前馈网络来表示,其结构如图1所示.

其中隐节点函数是小波函数,而输入层到隐层的权值和阈值分别对应小波的伸缩和平移参数,该参数可以通过对训练数据进行时频分析后获得,具体方法将在后面介绍.网络待确定的参数只有隐层到输出层的权值,且误差函数关于权值是一个凸函数,权值可以采用线性优化方法获得,不存在局部极小值问题,收敛速度较快.

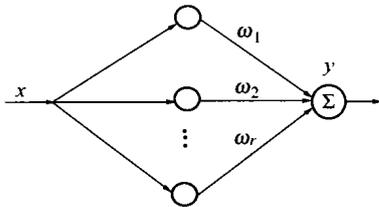


图1 基于框架的小波神经网络结构

Fig.1 Structure of wavelet neural networks based on frame

它是一张二维表格,每一行描述一个对象,每一列描述对象的一种属性.属性分为条件属性和决策属性.决策属性对于条件属性的依赖度是粗糙集理论中一个很重要的概念.

设属性 D 和属性 C 是属性集合 A 的两个子集,则 D 对 C 的依赖度定义为^[7]

$$k = \frac{|POS_C(D)|}{|U|}.$$

其中

$$POS_C(D) = \bigcup_{X \in U/D} C_*(X) \quad (0 \leq k \leq 1). \quad (3)$$

其中 $|*|$ 表示集合 $*$ 的基数或势(cardinality),对有限集合表示集合中所包含的元素的个数; $POS_C(D)$ 称为划分 U/D 关于 C 的正区; $I(x)$ 表示所有与 x 不可分辨的对象所组成的集合.

如果 $I(C) \subseteq I(D)$ 称属性 D 完全依赖于属性 C .由于条件属性和决策属性的每一对数据都可以看作是一条规则,因此实际计算时属性 D 对属性 C 的依赖度 k 可以用相容规则的个数除以总的规则个数求得.

3.2 结构优化方法基本原理 (Basic principle of structure optimization)

首先通过对训练数据和小波框架 $\{\Psi_{mn}(x)\}$ 进行时频分析,给出小波神经网络的初步构造.把网络输出作为决策表的决策属性,把网络隐层节点输出与权值的乘积作为条件属性.根据粗糙集理论分别计算决策属性对于条件属性的依赖度,它实际上反映了隐节点对网络输出贡献的大小,因此可以根据它每次删去对网络输出贡献最小的一个或几个隐节点.实际上本文的方法包含了文献[5]的仅根据权值大小删减隐节点的方法,这是因为较小的权值必然导致大量的条件属性值处于同一区间,从而使不相容规则数大大增加,依赖度下降.由于同时考虑了权值和隐节点输出,因此本文方法更为合理.

由于粗糙集理论只能处理离散的属性,因此有必要对连续属性进行量化.量化的任务是把连续属性的取值范围或取值区间划分为数目不太多的小区间,其中每个区间对应一个离散的符号,构成属性的领域.连续属性量化的方法有很多种,本文采用基于VDM(value difference metric)度量的划分法^[8].

3.3 具体步骤 (Specific steps)

① 确定信号 f 的时频集中区域.

由于对任一物理量的观测只能在有限的空间范围内进行,因此认为 $f(x)$ 在观测区间 $\Delta x = [x_{\min}, x_{\max}]$ 以外恒为0.虽然这种“截断”函数的频谱是无

3 基于粗糙集理论的小波神经网络结构优化设计方法 (Structure optimization of wavelet neural network using rough set theory)

3.1 粗糙集理论的相关知识 (Basic concepts of rough set theory)

3.1.1 粗糙集合的上逼近和下逼近 (Upper and lower approximation of rough set)

假定 U 和 A 是两个有限非空集合,其中 U 为论域, A 为属性的集合, B 为集合 A 的任意一个子集.

集合 X 关于 B 的下逼近定义为

$$B_*(X) = \{x \in U : B(x) \subseteq X\}.$$

它是由那些根据已有知识判断肯定属于的对象组成的集合.

集合 X 关于 B 的上逼近定义为

$$B^*(X) = \{x \in U : B(x) \cap X \neq \emptyset\}.$$

它是由那些根据已有知识判断可能属于 X 的对象组成的集合.

3.1.2 属性的依赖度 (Dependency of attributes)

粗糙集理论应用决策表来描述论域中的对象.

限扩展的,但总可以通过对训练数据的频谱分析估计出一个带宽区间 $\Delta\omega$, 使在此区间以外的能量密度可忽略不计,该区间可包括几个不连续的小区间. 那么 $Q_f = \Delta x \times \Delta\omega$ 就是信号 f 的时频集中区域.

② 确定小波框架 $\{\Psi_{mn}(x)\}$ 的时频集中区域 Q_{mn} .

③ 给出小波神经网络的初步构造.

计算使 Q_f 和 Q_{mn} 相交的所有 m, n 对, 记为集合 I :

$$I = \{(m, n): Q_{mn} \cap Q_f \neq \emptyset; m, n \in \mathbb{Z}\}. \quad (4)$$

取 $\Psi_{mn}(x), m, n \in I$ 作为隐层基函数构成小波神经网络.

④ 利用最小二乘法对网络权值进行训练, 根据 3.2 节中的方法构造决策表.

⑤ 利用粗糙集理论中有关依赖度的概念, 每次去除依赖度最小的一个或数个隐层节点; 重新对网络进行训练后判断性能指标是否满足要求; 如果仍满足, 则转回 ④, 否则前一次所得小波神经网络为最终结果.

4 仿真研究 (Simulation study)

仿真研究采用与文献[1]相同的例子, 利用小波神经网络来逼近如式(5)的分段函数

$$f(x) = \begin{cases} -2.186x - 12.864 & (-10 \leq x < -2), \\ 4.246x & (-2 \leq x < 0), \\ 10e^{-0.05x-0.5} & (0 \leq x \leq 10). \end{cases} \quad (5)$$

x 在区间 $[-10, 10]$ 上均匀选取 200 点作为训练数据. 取 Ψ 为 Mexico 小波, 即

$$\Psi(x) = (1 - x^2)\exp(-x^2/2), x \in \mathbb{R}. \quad (6)$$

如果取伸缩参数 $a = 2$, 平移参数 $b = 0.5$, 则 $\{\Psi_{mn}\}$ 构成 $L^2(\mathbb{R})$ 中的小波框架^[9]. 对训练数据进行时频分析得到具有 52 个隐节点的小波神经网络的初步构造. 在此基础上应用第 3 节基于粗糙集理论的结构优化设计方法, 最终得到的小波神经网络隐节点个数为 21 个. 为便于比较, 采用与文献[1]相同的性能指标

$$J = \sqrt{\frac{\sum_{l=1}^L (\hat{y}_l - y_l^d)^2}{\sum_{l=1}^L (y_l^d - \bar{y}_l)^2}}. \quad (7)$$

其中 $\bar{y} = \frac{1}{L} \sum_{l=1}^L y_l^d$, y_l^d 为函数实际输出, \hat{y}_l 为网络输出.

函数学习的最终结果如图 2 所示. 与文献[1]方法结果的比较见表 1.

从图 2 和表 1 可以看出: 本文基于粗糙集理论的小波神经网络的结构优化设计方法是成功的. 一方面它减轻了用小波框架作为隐层基函数所带来的冗余性, 使网络隐节点的个数大大减少; 另一方面还使网络的性能有所提高. 简单和易于理解也是该方法的优点之一.

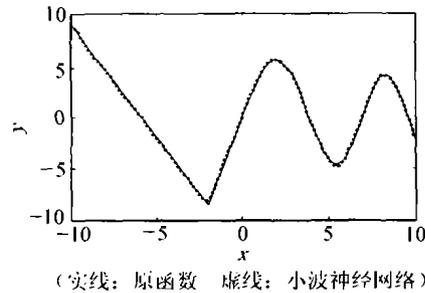


图 2 原函数与小波神经网络输出的比较
Fig. 2 Comparison of outputs between original function and the WNN

表 1 与文献[1]方法结果的比较
Table 1 Comparison of our results with Reference[1]

方法	未知参数个数	性能指标 J
本文	21	0.0253
文献[1]	22	0.05057

5 结论 (Conclusion)

针对基于框架的小波神经网络存在的网络结构冗余性较大的问题, 本文提出一种基于粗糙集理论的网络结构优化设计方法. 首先通过对训练数据和小波框架 $\{\Psi_{mn}(x)\}$ 进行时频分析, 确定小波神经网络的初步构造; 在此基础上, 根据网络输出对隐层节点依赖度的大小去除冗余的隐层节点, 达到优化网络结构的目的. 与以往的方法^[4,5]相比, 本文方法显得简单而有效.

参考文献 (References):

- [1] ZHANG Qinghua, BENVENISTE A. Wavelet networks [J]. *IEEE Trans on Neural Networks*, 1992, 3(6): 889 - 898.
- [2] HO D W C, ZHANG Pingan. Fuzzy wavelet networks for function learning [J]. *IEEE Trans on Fuzzy Systems*, 2001, 9(1): 200 - 211.
- [3] HUANG Dexian, WANG Jingchun, JIN Yihui. Application research of wavelet neural networks in process control [J]. *J of Tsinghua University*, 1999, 39(1): 91 - 94.
- [4] 张茁生, 刘贵中, 刘峰. 一种自适应小波神经网络的构造及其学习算法 [J]. *中国科学 (E 辑)*, 2001, 31(2): 172 - 181.

(ZHANG Zhuosheng, LIU Guizhong, LIU Feng. A construction and

(下转第 337 页)

- ence, 1990, 1(1):3-26.
- [7] SHREEDHAR M, VARGHESE G. Efficient fair queuing using deficit roundrobin [J]. *IEEE/ACM Trans on Networking*, 1996, 4(3): 375-385.
- [8] MISHRA P, KANAKIA H, TRIPATHI S. On hop-by-hop rate-based congestion control [A]. *IEEE/ACM Trans on Networking*, 1996, 4(2):224-239.
- [9] FLOGD S, JACOBSON V. Random early detection gateways for congestion avoidance [J]. *IEEE Trans/ACM on Networking*, 1993, 1(4):397-413.
- [10] ROMANOW A, FLOGD S. Dynamics of TCP traffic over ATM networks [J]. *IEEE J on Communications*, 1995, 13(4):633-641.
- [11] ZHANG Lixia, DEEYING S, ESTRIN D, et al. A new resource reservation protocol [J]. *IEEE Network*, 1993, 7(5):8-18.
- [12] 刘锦德,刘后铭,周明天,等.计算机网络大全[M].北京:电子工业出版社,1997.
(LIU Jinde, LIU Houming, ZHOU Mingtian, et al. *Entirety of Computer Network* [M]. Beijing: Electronic Industry Press, 1997.)
- [13] 王立新.自适应模糊系统与控制[M].北京:国防工业出版社, 1995.
(WANG Lixin. *Adaptive Fuzzy System and Control* [M]. Beijing: National Defense Industry Press, 1995.)

作者简介:

杨洪勇 (1967—),男,博士生,副教授.主要研究领域为自动控制理论,网络通信理论及应用. E-mail: yhy9919@163.com;

武玉强 (1962—),男,博士,教授.目前主要研究领域为变结构控制,非线性系统理论,网络通信理论及应用. E-mail: wyq@qfnu.edu.cn.

(上接第331页)

- learning algorithm on adaptive wavelet neural networks [J]. *Science in China (Series E)*, 2001, 31(2):172-181.)
- [5] ZHANG Qinghua. Using wavelet network in nonparametric estimation [J]. *IEEE Trans on Neural Networks*, 1997, 8(2):227-236.
- [6] 李永敏,朱善君,陈湘晦.根据粗糙集理论进行BP网络设计的研究[J].系统工程理论与实践,1999,19(4):62-69.
(LI Yongmin, ZHU Shanjun, CHEN Xianghui. A rough set approach to BP neural network designing [J]. *System Engineering—Theory & Practice*, 1999, 19(4):62-69.)
- [7] PAWLAK Z. Why rough sets [A]. *Proc of the fifth IEEE Int Conf on Fuzzy Systems* [C]. Piscataway, NJ: IEEE Press, 1996:738-743.
- [8] 张飞弓,叶东毅.粗糙集连续属性量化算法的灵敏度分析[J].福州大学学报,2000,28(3):28-30.
(ZHANG Feigong, YE Dongyi. Sensitivity analysis of quantifying algorithm for continuous attributes of rough set [J]. *J of Fuzhou University*, 2000, 28(3):28-30.)
- [9] DAUBECHIES I. The wavelet transform, time-frequency localization and signal analysis [J]. *IEEE Trans on Information Theory*, 1990, 36(5):961-1005.

作者简介:

李益国 (1973—),男,讲师,博士.主要研究方向为智能控制理论在热工过程控制中的应用. E-mail: li-yiguo@sohu.com;

沈炯 (1957—),男,教授,博士生导师.主要研究方向为先进控制理论在热工过程控制中的应用;

吕震中 (1941—),男,教授,博士生导师.主要研究方向为热工过程控制及微机应用.