

多层前向小世界神经网络的逼近与容错性能

李小虎¹, 杜海峰², 张进华¹, 王孙安¹

(1. 西安交通大学机械工程学院, 710049, 西安; 2. 西安交通大学公共管理与复杂性科学研究中心, 710049, 西安)

摘要: 基于 Watts-Strogatz 网络模型的构造思想, 对多层前向神经网络中的规则连接依重连概率进行重连, 构建了一种多层前向小世界神经网络模型. 对该网络模型进行简要的数学描述, 并以函数逼近和网络容错仿真考察了构建的小世界神经网络的性能. 结果表明, 与规则或随机连接的网络相比, 当重连概率处于 0.1~0.2 时, 小世界神经网络具有更优的逼近性能, 且当网络学习速率参数在 0.1~0.3 之间时, 对小世界神经网络的逼近性能影响较小. 此外, 当网络权值故障率小于 30% 时, 重连概率不大于 0.8 的小世界神经网络与规则网络同样具有较优的容错性能, 而当故障率大于 40% 时, 重连概率较大的小世界神经网络和随机连接的神经网络的容错性能要明显优于规则网络.

关键词: 小世界神经网络; 复杂网络; 函数逼近

中图分类号: TP183 **文献标志码:** A **文章编号:** 0253-987X(2010)07-0059-05

Approximation and Fault-Tolerance Performances of Multilayer Feedforward Small World Neural Network

LI Xiaohu¹, DU Haifeng², ZHANG Jinhua¹, WANG Sun'an¹

(1. School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an 710049, China; 2. Center for Administration and Complexity Science, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: Based on the idea of Watts-Strogatz (W-S) network model, a multilayer feedforward small world neural network, which relies heavily on the rewiring probability, is constructed by reconnecting the regular links. The network model is described mathematically, and the function approximation and network fault-tolerance simulations are employed to investigate the performances of small world neural network. The results show that the small world neural network is endowed with better approximation performance compared with the regular or random connecting network; and when the learning rate parameter of network gets between 0.1 to 0.3, it exerts less impact on the approximation performance of small world neural network. In addition, when the network weight failure is less than 30%, small world neural network with rewiring probability less than 0.8 and regular network have the same better performance on fault-tolerance. When failure rate is more than 40%, small world neural network with high rewiring probability and random connecting neural network have distinctly better performance on fault-tolerance than the regular neural network has.

Keywords: small world neural network; complex network; function approximation

大脑神经网络是由大量神经元通过神经纤维连接而成, 虽然单个神经元的结构和功能相对简单, 但用某种形式相互连接则可形成高度复杂的网络^[1]. 大脑网络的功能区定位研究表明, 大脑的结构与功

能之间密切相关,各个功能区在完成大脑功能活动时相互协作^[2],可见大脑神经元的连接结构为各种协作提供了可能.人工神经网络模型是对大脑神经网络结构和功能的简单模拟,网络结构的连接模式常采用前向、反馈、单层和多层等多种规则连接形式^[3],对这些人工神经网络模型的研究比较多^[4-6],而对有关非规则连接结构的人工神经网络模型的研究相对较少.如何设计非规则连接模式的网络模型是一个首要的问题,网络连接模式设计的目的是为了获得最优的网络性能,且能较为真实地反映大脑神经网络的结构特征.

Watts 和 Strogatz^[7]指出现实网络在结构上多为从规则连接到随机连接的中间网络模型,并定义为小世界网络(简称 W-S 模型).近 10 年来,大量关于复杂网络的研究表明,现实网络如 WWW 网、疾病传播网、社会关系网、食物链网、新陈代谢网等多为小世界网络^[8-9],同时许多学者在对生物神经网络进行研究时也发现其具有小世界特征^[10-11].因此,在结构上将人工神经网络模型构建成为具有小世界特征的网络是对生物神经网络较为真实的模拟. Simard 等^[12]通过对该网络中的规则连接进行重连来构建小世界神经网络,利用其对二进制样本学习时发现小世界神经网络比规则网络和随机网络具有更快的学习速度和更小的误差,但所采用的重连边数量的确定却缺乏对网络结构从规则到完全随机的变化过程的反映.

本文首先参考 W-S 模型的构建过程,提出一种多层前向小世界神经网络模型的构建算法,然后对模型进行简要的数学描述,并将所构建的多层前向小世界神经网络与规则网络的特性进行对比,探讨了小世界神经网络的逼近性能.最后在权值连接出现故障时,对比了小世界神经网络与规则、随机连接的神经网络的容错性能.

1 多层前向小世界神经网络的构建

小世界现象揭示了网络在具有高度聚集的局部

连接的同时,也包含一些有助于产生短路径的随机长连接^[13-14].文献[12]参考 W-S 模型的构造过程,对多层前向规则神经网络中的以连接依据确定的连接边数重连,构建了多层前向小世界神经网络,而这种并未按重连概率 p 所得到的重构网络具有一定的狭义性,不能完全反映小世界网络的构造过程.考虑到 W-S 模型中 p 的大小会直接影响所生成网络的结构和特性,又由于输入输出层的神经元个数由问题决定,因此本文提出一种根据重连概率对多层前向神经网络的隐层进行随机重连,构造广义多层前向小世界神经网络模型的算法,步骤如下.

步骤 1 初始化:设定网络隐层数 L ,隐层神经元个数 $n_l (l=1,2)$ 和 p .

步骤 2 生成网络连接规则的多层前向神经网络,网络中第 l 层的第 i 个神经元 v_{li} 只与相邻层的神经元集 V_{l-1} 和 V_{l+1} 中的神经元相连,且所有连接均为前向连接,各层内神经元之间无连接.

步骤 3 重连迭代次数 $m \leftarrow 1$.

步骤 4 以 p 断开网络中第 m 层中的神经元 v_{mi} 指向第 $m+1$ 层神经元 $v_{(m+1)j}$ 的连接,并重新随机在第 $m+2 \sim L$ 层中选择神经元 v_j 进行重新连接,但新连接与重连和自环无关.

步骤 5 $m \leftarrow m+1$.

步骤 6 算法终止条件判断:如果 $m \leq L-1$,则返回步骤 4,否则算法终止.

由于网络中第 $L-1$ 层的规则连接断开后无法产生新的长连接,因此不对网络中最后两隐层神经元间的连接进行随机重连.图 1 给出了 $L=6, n_l=5$ 的小世界神经网络的构建过程.由图 1a 可见,网络保持了完全规则连接的模式,而图 1c 表示网络中所有的原连接(除最后两层间的连接)均发生重连,此时网络的连接结构具有完全随机的特性.当 $0 < p < 1$ 时,网络依 p 产生跨层长连接,这是一种在结构上介于完全规则和完全随机之间的网络模型(见图 1b),这种连接结构是从有序到无序的过渡形式,符合小世界网络的定义^[7].此外,当网络规模和 p 都

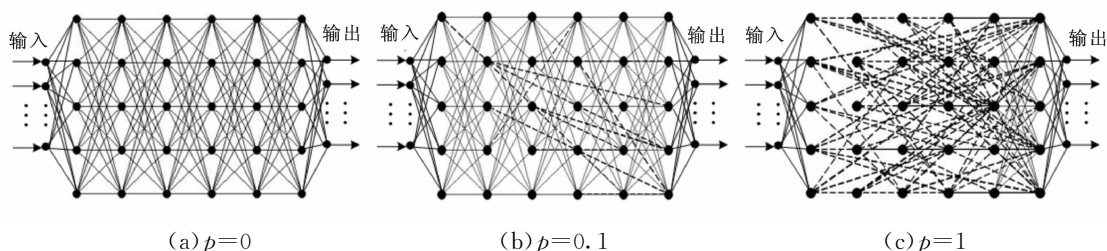


图 1 多层前向小世界神经网络模型的构建过程

较小时,网络中引入的随机重连数也较少。

2 多层前向小世界神经网络模型描述

神经网络可以用图来描述,多层前向小世界神经网络中各个神经元为图的节点,神经元之间的连接为图的边^[15]. 小世界神经网络中的隐层神经元可用节点集合 $V = \{V_1, \dots, V_l, \dots, V_L\}$ 来表示,其中 $l \in \{1, 2, \dots, L\}$, $V_l = \{v_i | i=1, 2, \dots, n_l\} \subset V$ 为第 l 隐层的节点子集. 网络的隐层连接矩阵为

$$\mathbf{W} = \{\mathbf{W}_1, \dots, \mathbf{W}_l, \dots, \mathbf{W}_{L-1}\} \quad (1)$$

式中: $\mathbf{W}_l (l=1, 2, \dots, L-1)$ 为 $V_l \rightarrow V_{l+1}$ 层的连接子矩阵. $\mathbf{w}_{li} \subset \mathbf{W}_l$ 为 $V_l \rightarrow v_{(l+1)i}$ 的连接向量, $i=1, 2, \dots, n_l$, $\mathbf{w}_{li} \in \mathbf{R}$ 为节点 $v_i \rightarrow v_{(l+1)i}$ 的连接权. 如果隐层中第 l 层的神经元 v_i 与第 $l+1$ 层的神经元 $v_{(l+1)j}$ 之间有连接, $\mathbf{w}_{li} \neq 0$; 反之, $\mathbf{w}_{li} = 0$, 且不考虑神经元的自连接. 因此, 对于规则网络, 隐层连接矩阵可表示为

$$\mathbf{W} = \begin{matrix} & \begin{matrix} V_1 & V_2 & V_3 & \cdots & V_{L-1} & V_L \end{matrix} \\ \begin{matrix} V_1 \\ V_2 \\ \vdots \\ V_{L-1} \\ V_L \end{matrix} & \begin{bmatrix} 0 & \mathbf{W}_1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \mathbf{W}_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & \mathbf{W}_{L-1} \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \end{matrix} \quad (2)$$

对于多层前向小世界神经网络, 由于重连将使网络的隐层连接矩阵变为

$$\mathbf{W}' = \begin{matrix} & \begin{matrix} V_1 & V_2 & V_3 & \cdots & V_{L-1} & V_L \end{matrix} \\ \begin{matrix} V_1 \\ V_2 \\ \vdots \\ V_{L-1} \\ V_L \end{matrix} & \begin{bmatrix} 0 & \mathbf{W}'_1 & \mathbf{B}_{13} & \cdots & \mathbf{B}_{1(L-1)} & \mathbf{B}_{1L} \\ 0 & 0 & \mathbf{W}'_2 & \cdots & \mathbf{B}_{2(L-1)} & \mathbf{B}_{2L} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & \mathbf{W}'_{L-1} \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \end{matrix} \quad (3)$$

因此网络中最后 2 层间无重连. $\mathbf{W}'_{L-1} = \mathbf{W}_{L-1}$, $\mathbf{W}'_1 = (\mathbf{w}'_{1ij})_{n_1 \times n_1}$ 为 $\mathbf{W}_1$ 重连后的矩阵, 该矩阵中的元素可表示为

$$\mathbf{w}'_{1ij} = \begin{cases} \mathbf{w}_{1ij}, & \text{rand}(i, j) \geq p \\ 0, & \text{rand}(i, j) < p \end{cases} \quad (4)$$

其他重连矩阵的定义和取值与 $\mathbf{W}_1$ 类似. $\mathbf{B}_{13} = (\mathbf{b}_{13}^{13})_{n_1 \times n_1}$ 表示 $V_1 \rightarrow V_3$ 的重连子矩阵, 若第 l 层连向第 l' 层的重连子矩阵为 $\mathbf{B}_{ll'} = (\mathbf{b}_{ll'}^{ll'})_{n_l \times n_{l'}}$, 其中 $l \in \{1, 2, \dots, L-2\}$, $l' \in \{3, 4, \dots, L\}$, 该重连子矩阵产生的概率为 $1/(L-l-1)$, 该概率即为第 $l+1$ 层至第 L 层中的任意层被选中的概率. 选中层中的第 n_2 个神经元在同层神经元中被选中的概率为 $1/n_l$, 则

重连子矩阵中的元素可表示为

$$b_{n_1 n_2}^{ll'} = \begin{cases} \text{rand}, & \text{rand} < p/[(L-l-1)n_l] \\ 0, & \text{rand} \geq p/[(L-l-1)n_l] \end{cases} \quad (5)$$

由 $\mathbf{W}'$ 可以看出, 网络中第 l 层的节点依概率 $1-p$ 连接至第 $l+1$ 层的神经元, 依 p 产生重连至后续层神经元. 若网络中层数和神经元数较多且 p 较小时, 则 $\mathbf{W}'$ 为稀疏矩阵.

在函数逼近测试中, 网络的输入向量 $\mathbf{X}$ 由逼近函数的自变量所组成, 输出向量 $\mathbf{Y}$ 由多层前向小世界神经网络输出层神经元的输出组成, 每个神经元的激活函数采用非线性双极性 Sigmoid 函数. 网络学习算法采用标准的误差反传算法进行权值修正, 使逼近误差趋近于 0, 实现输入到输出的映射: $\mathbf{X} \rightarrow \mathbf{Y}$, 从而实现复杂问题的网络表达, 其中学习速率为 η . 本文根据逼近函数输入输出变量的维数, 采用

的网络结构为 $p_1 \times \overbrace{n_l \times \cdots \times n_l}^k \times q_1$, 其中 p_1 为逼近函数输入变量的维数, q_1 为输出变量的维数. 构建小世界网络时, 对网络中的隐层部分 $\overbrace{n_l \times \cdots \times n_l}^k$ 进行随机重连, 仿真试验不进行离线训练, 所有结果均为直接测试结果.

3 小世界神经网络的逼近及容错性能

3.1 逼近性能试验

为了考察多层前向小世界网络的逼近性能, 采用二维草帽函数作为逼近函数, 其表达式为

$$y = \frac{\sin(\pi(x_1^2 + x_2^2)^{1/2})}{\pi(x_1^2 + x_2^2)^{1/2}}, \quad x_1, x_2 \in [-5, 5] \quad (6)$$

在该函数的给定自变量区间均匀选择 2 500 个点, 分别在 $p=0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0$ 时, 考察各个网络对上述草帽函数的逼近性能. 由于逼近函数有 2 个输入变量, 因此选择的网络结构为 $2 \times \overbrace{n_l \times \cdots \times n_l}^k \times 1$, 其中 $n_l=8, L=10$. 采用标准的误差反传算法进行权值更新, $\eta=0.2$, 算法终止条件为迭代 5 000 次. 为了消除随机性, 每个 p 下均迭代 30 次后取平均值. 从表 1 中可以看出, 规则和完全随机网络的逼近误差 e_a 都较大, 当 p 增大时, e_a 也在增大, 这是由于 p 增大使网络中出现的重连越来越多, 其结构特性与完全随机连接的网络比较接近. 最优误差值出现在 $p=0.1 \sim 0.2$ 附近, 表明只需通过少数重连, 便可提高网络的函数逼近

表1 不同重连概率率下的平均逼近误差

p	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
e_a	1.117 7	0.888 6	0.669 3	0.977 4	0.903 5	1.018 1	0.960 5	1.026 9	1.044 6	1.135 7	1.167 2

性能.

为了进一步说明各个 p 下 e_a 的变化情况,图2给出了 $p=0.0,0.1,0.2,1.0$ 时, e_a 曲线的变化情况.由图2可知, $p=0.0$ 和 $p=1.0$ 的 e_a 最大,而 $p=0.1$ 和 $p=0.2$ 的 e_a 相对较小,其中 $p=0.2$ 时 e_a 最小,此时小世界神经网络的收敛速度也较快.由图3可知,当 $p=0.2$ 时,网络逼近输出的效果较好.

3.2 参数影响试验

由于小世界神经网络与多层规则神经网络一样,都采用了标准的BP算法,因此为了验证 η 对小世界神经网络的性能影响,分别在 $p=0.1,0.3$ 时采用二维草帽函数进行逼近测试,相关参数与逼近性能试验一致.图4给出了 η 从 $0.05\sim0.45$ 变化时的测试结果,当 η 在 $0.1\sim0.3$ 之间时,所得 e_a 变化不大,可见此时小世界神经网络逼近性能受参数影响较小,而当 η 小于 0.1 或大于 0.3 时, e_a 显著增大,表明 η 是一个重要的参数,偏大偏小都不利于小世界神经网络获得较好的逼近性能,这与多层前向规则神经网络的特性基本一致.

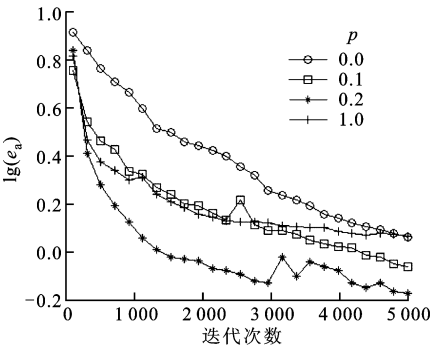


图2 逼近误差随重连概率的变化趋势

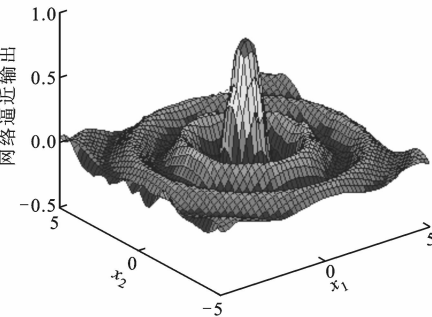


图3 网络逼近输出($p=0.2$)

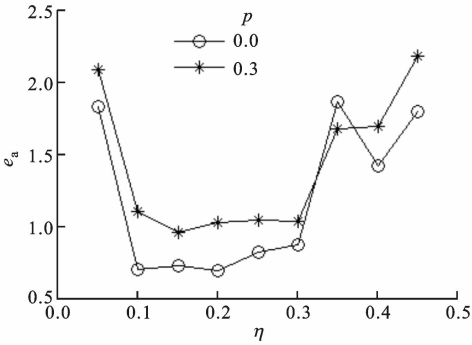


图4 学习速率取值对 e_a 的影响

3.3 容错性能试验

为了验证小世界神经网络的容错特性,选择二维草帽函数作为容错性能测试函数,具体形式如下

$$y = \frac{2}{3^{1/2}} \pi^{-0.25(1-x^2)} e^{-x^2/2}, x \in [-5,5] \tag{7}$$

在给定的自变量区间,均匀选择 $N=101$ 个数据点逼近该函数.神经网络采用单输入单输出网络,具体形式为 $1 \times \overbrace{n_l \times \cdots \times n_l}^L \times 1$,其中 $n_l=6, L=10$.算法的其他参数与逼近性能测试相同,网络容错性能测试主要模拟神经元之间连接断开的情况.

图5给出了容错性能测试的结果,图5a是断开率 σ 相对较低的情况, σ 分别为 $10\%\sim30\%$.由图中曲线变化的情况可以看出,当网络连接故障率较低时,规则网络和 p 不大于 0.8 的网络基本上同样具有较好的容错能力,其 e_a 都较小,而当 p 大于 0.8 时, e_a 会迅速变大,表明随机连接较多的网络容错能力变差.图5b给出了 σ 较大的几种情况, σ 分别为 $40\%\sim80\%$,图中曲线显示当网络处于规则连接时, e_a 较大,表明网络此时容错能力较差.当 p 增大时, e_a 开始减小,且在 $p=1.0$ 时, e_a 比较小,这表明当网络的 σ 较大时, p 较大的小世界神经网络和随机连接的神经网络都具有比规则网络更好的容错能力.

4 结论与展望

本文借鉴复杂网络的研究成果,根据重连概率构建了一种在结构上处于规则和随机之间的多层前向小世界神经网络模型,并对其性能进行了研究.函数逼近性能仿真测试结果表明,只需增加少量随机

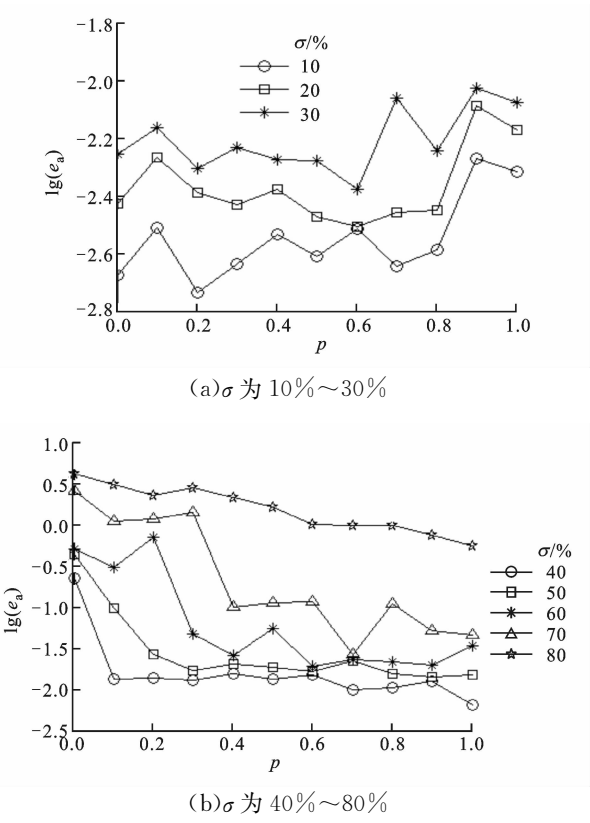


图 5 不同断开率下的逼近误差

长连接,即重连概率在 0.1~0.2 左右时,小世界神经网络比规则和随机连接网络具有更优的逼近性能.当学习速率处于 0.1~0.3 之间,小世界神经网络的逼近性能大致相同,表明此时网络具有较好的稳定性.此外,网络的容错性能对比试验结果表明,当网络断开率大于 40%时,多层前向小世界神经网络和随机神经网络的容错性能要优于规则网络.近年来,有关复杂网络的研究成果日新月异,下一步的研究工作是将这些成果用于构建新的人工神经网络模型.

参考文献:

[1] 李翔. 从复杂到有序:神经网络智能控制理论新进展 [M]. 上海:上海交通大学出版社, 2006.

[2] SPORNS O, HONEY C J. Small worlds inside big brains [J]. Proceedings of the National Academy of Sciences of the United States of America, 2006, 103 (51):19219-19220.

[3] 袁著祉, 陈增强, 李翔. 联接主义智能控制综述 [J]. 自动化学报, 2002, 28(1):38-59.

YUAN Zhuzhi, CHEN Zengqiang, LI Xiang. Connectionism intelligent control: a survey [J]. Acta Automatica Sinica, 2002, 28(1):38-59.

[4] XIA Y S, WANG J. A general methodology for designing globally convergent optimization neural networks

[J]. IEEE Trans on Neural Networks, 1998, 9(6): 1331-1343.

[5] MURATA N. Network information criterion determining the number of hidden units for an artificial neural network model [J]. IEEE Trans on Neural Networks, 1994, 5(6):865-872.

[6] HORNIK K, STINCHCOMBE M, WHITE H. Multilayer feedforward networks are universal approximators [J]. Neural Networks, 1989, 2(2):359-366.

[7] WATTS D J, STROGATZ S H. Collective dynamics of small-world networks [J]. Nature, 1998, 393(4): 440-442.

[8] ALBERT R, BARABA'SI A L. Statistical mechanics of complex networks [J]. Reviews Modern Physics, 2002, 74(1):47-97.

[9] 杜海峰, 庄健, 张进华, 等. 用于函数优化的小世界优化算法 [J]. 西安交通大学学报, 2005, 39(9):1011-1015.

DU Haifeng, ZHUANG Jian, ZHANG Jinhua, et al. Small-world phenomenon for function optimization [J]. Journal of Xi'an Jiaotong University, 2005, 39 (9):1011-1015.

[10] SPORNS O, TONONI G, EDELMAN G M. Connectivity and complexity: the relationship between neuroanatomy and brain dynamics [J]. Neural Networks, 2000, 13(8):909-922.

[11] BASSETT D S, BULLMORE E. Small-world brain networks [J]. Neuroscientist, 2006, 12(6):512-523.

[12] SIMARD D, NADEAU L, KROGER H. Faster learning in small-world neural networks [J]. Physics Letters: A, 2005, 336(1):8-15.

[13] 杜海峰, 李树苗, MARCUS W F, 等. 基于先验知识和模块性的网络社区结构探测算法 [J]. 西安交通大学学报, 2007, 41(6):750-754.

DU Haifeng, LI Shuzhuo, MARCUS W F, et al. Detecting algorithm based on prior knowledge and modularity for networked community structure [J]. Journal of Xi'an Jiaotong University, 2007, 41(6): 750-754.

[14] KLEINBERG J. The small-world phenomenon and decentralized search [J]. SIAM News, 2004, 37(3):1-3.

[15] 阮晓钢. 人工智能之路神经计算科学 [M]. 北京:国防工业出版社, 2006.

[16] 阎平凡. 对多层前向神经网络研究的几点看法 [J]. 自动化学报, 1997, 23(1):129-135.

YAN Pingfan. Some views on the research of multilayer feed forward neural networks [J]. Acta Automatica Sinica, 1997, 23(1):129-135.

(编辑 管咏梅)