

# 用于混沌时间序列预测的多簇回响状态网络<sup>\*</sup>

宋青松<sup>†</sup> 冯祖仁 李人厚

(西安交通大学系统工程研究所, 西安 710049)

(2008 年 1 月 7 日收到, 2008 年 12 月 9 日收到修改稿)

研究了混沌时间序列预测问题, 提出了一种由五元生长因子组调控的类皮层神经网络模型, 即多簇回响状态网络模型(MCESN). 研究表明该生长因子组能够有效决定模型的拓扑性质, 同时具备小世界和无标度等复杂网络特征的 MCESN 能够获得较优的预测结果. 通过 Monte Carlo 仿真实验表明, 该模型不仅训练算法简单, 而且与常规回响状态网络比较, 预测结果的精度更高、标准差更小.

关键词: 混沌时间序列预测, 回响状态网络, 复杂网络,  $\Omega$  复杂性

PACC: 8710, 0547

## 1. 引言

时间序列的预测在多个领域具有重要的意义<sup>[1-3]</sup>. 根据构成时间序列的数据点之间是否存在相关, 时间序列分为非静态时间序列和静态时间序列. 非静态时间序列预测由于内在的短时段记忆 (STM) 需求, 通常难于静态时间序列预测, 比如一类典型的非静态时间序列, 即混沌时间序列的预测, 往往是一个棘手的问题<sup>[4]</sup>. 神经网络方法是用于混沌时间序列预测的常见方法之一, 实现方式通常分为两类. 一类是基于时间窗口的方法. 这类方法首先定义一个固定长度为  $M\tau$  的时间窗口  $tw(t) = [u(t), u(t-\tau), \dots, u(t-(M-1)\tau)]$ , 然后基于该时间窗口完成预测  $y(t) = f(tw(t))$ , 比如延时神经网络方法<sup>[5]</sup>. 其本质是将时间域上的非静态转化为空间域上的静态. 另一类方法是在前向感知机中增加回归连接, 生成回归神经网络 (RNNs). ‘环路’的存在导致网络活性在网络中被自我维持一段时间, 即 STM 现象, 比如 Jordan 网络, 回归多层感知机 (RMLPs)<sup>[6]</sup> 等. 在这两类方法中, 时间窗口宽度  $M\tau$ , 回归连接个数及环路发生位置都难以确定, 通常采用 tried-and-error 方法确定, 并且一旦确定, 难以根据任务动态改变, 从而缺少适应性. 此外, 基于误差梯度的神

经网络训练算法具有陷入局部极小, 甚至不收敛等缺陷. 正是由于这些不足, 导致神经网络方法特别是 RNNs 没有得到广泛使用.

回响状态网络<sup>[7]</sup> (echo state network, ESN) 和液体状态机<sup>[8]</sup> (liquid state machine, LSM) 是一类新型大规模回归神经网络模型. 使用 ESN 对 MG 混沌时间序列进行 84 步前向预测, 得到归一化均方误差 (NRMSE) 为  $10^{-4.2[9]}$ . 相对于目前常见的神经网络方法得到的约为  $10^{-1.2}$ — $10^{-1.7}$  的 NRMSE<sup>[10]</sup>, ESN 预测精度得到几乎两个数量级的质的改善, 并且简单的线性回归即可胜任该训练任务. 文献 [11—13] 也构建出特殊结构的 ESN, 对相应问题求得更满意解. 但是, 常规 ESN 方法存在一个缺陷: 同一问题的 Monte Carlo 仿真实验结果方差较大, 不具备平稳分布特征<sup>[14]</sup>.

针对这个缺陷, 在 Kaiser 等人提出的多簇网络生长算法<sup>[15, 16]</sup>基础上, 本文实现由 5 个网络生长参数刻画的一类多簇 ESN. 期望该网络对于混沌时间序列预测问题, 不仅继承常规 ESN 方法中训练简单的优势, 而且保证预测结果具备小误差, 小方差优势, 从而为设计高效 ESN 提供一个有效方法.

## 2. 回响状态网络

ESN 体系结构如图 1. 由相对独立的三个部分,

<sup>\*</sup> 高等学校博士学科点专项科研基金 (批准号: 20050698032), 国家重点基础研究发展计划 (973 项目) (批准号: 2007CB311006), 国家高技术研究发展计划 (863 计划) (批准号: 2006AA04Z222) 和国家自然科学基金 (批准号: 60875043) 资助的课题.

<sup>†</sup> E-mail: qssong@sei.xjtu.edu.cn

即前端输入、中间部分的动态库(DR)和后端读取器组成。DR负责对输入激励信号的表达,往往是由众多神经元个体随机地稀疏连接而构成的大规模回归网络。该DR将输入激励映射到一个充分高维的空间,保证该输入激励在该高维空间中得到充分表达。后端读取器负责从大规模DR中萃取信息,通过DR与该读取器之间的连接权重线性叠加,产生期望输出。ESN的状态方程和输出方程如(1)式和(2)式:

$$X(n+1) = f_1(W^{\text{in}}u(n+1) + WX(n) + W^{\text{back}}Y(n)), \quad (1)$$

$$Y(n+1) = f_2(W^{\text{out}}X(n+1)), \quad (2)$$

其中,  $u$ ,  $X$  和  $Y$  分别表示输入向量、DR状态向量和输出向量。 $W^{\text{in}}$ ,  $W^{\text{out}}$ ,  $W$  和  $W^{\text{back}}$  分别是输入和输出连接权重矩阵、DR连接权重矩阵、输出反馈连接矩阵。 $f_1$  和  $f_2$  分别是DR中个体神经元和读取器单元的活性函数,该函数根据任务需要,可以选择为线性函数,也可以选择为非线性函数。本文选择双曲正切函数。

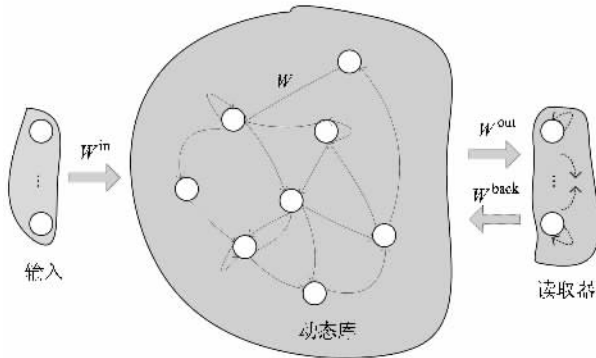


图1 ESN的体系结构示意图

Jaeger的研究<sup>[7]</sup>中,DR规模  $N$  (即构成库的神经元个数),库邻接矩阵的谱半径  $\rho(W)$  以及库连通密度(即网络中任意两个神经元个体之间存在的连接个数和网络中可能存在连接个数的比率)是刻画DR拓扑特征的三个主要参数。相对于规模通常在3—20个神经元的常见的人工神经网络,ESN往往取40—1000个神经元的规模。 $\rho(W)$ 取值小于1,以保证库回响状态属性和库动力学稳定性。库连通密度通常取0.01数量级,以保证库状态响应的多样性。根据这些先验约束,随机产生  $W$ ,完成ESN初始化。训练过程中,根据其状态更新方程式(1)将理想输出  $Y$  作为教师信号通过输出反馈连接权重  $W^{\text{back}}$  送入库,在该教师信号和输入激励信号  $u$  共同作用

下,从某个时刻  $T_0$  开始,收集DR状态  $X(n)$  和理想输出  $Y(n)$ ,  $n = T_0, \dots, T_K$ , 分别得到收集状态矩阵  $[ \cdot ]_{N,C}$  和理想输出矩阵  $[ \cdot ]_Y$ ,  $C = (T_K - T_0 + 1)$ 。在最小二乘意义下,可以得到

$$W^{\text{out}} = (f_2^{-1}([ \cdot ]_Y) \mathbb{I} - [ \cdot ]_{N,C} X [ \cdot ]_{N,C} [ \cdot ]_{N,C}^{-1})^{-1}. \quad (3)$$

从而完成ESN系统训练过程。将  $W^{\text{out}}$  写入输出权重连接矩阵,完成系统黑箱建模。之后对该模型进行测试和使用,以得到对该模型建模质量的评价。

可见ESN与支持向量机(SVM)等基于核的方法机理类似,都将输入映射到一个高维空间,读取机构从该高维空间中抽取信息,产生期望输出。但是由于DR中存在大规模稀疏回归连接;回响状态'具备STM特性。ESN的训练思想不同于通常人工神经网络中所有连接权重  $\{W^{\text{in}}, W, W^{\text{back}}, W^{\text{out}}\}$  都需要根据误差梯度进行学习调整的思路,整个ESN中,仅  $W^{\text{out}}$  需要学习调整;并且简单的线性回归就可胜任该训练任务。ESN这种将'信息表达'与'权重学习'独立实现的方法,回避了基于误差梯度的学习算法的难点,比如收敛速度慢,容易陷入局部极小<sup>[17]</sup>等。

### 3. 多簇回响状态网络

Sporns, Kaiser等<sup>[18,19]</sup>借助复杂网络分析方法,针对猫和短尾猿皮层连通网络,以及 *C. elegans* 神经网络,从全局的角度,寻找具有容错、信息快速处理以及平衡网络活性等性能的复杂大脑网络的拓扑特征和组织方式,以及这些拓扑特征和组织方式与通过脑功能成像或者EEG测量的功能表达之间的联系。他们发现空间上分布的大脑皮层网络并不是完全连通的,所有可能的连接中仅存在大约30%连接,并且发现存在高度连通的神经元节点集合,即神经元节点簇。这些节点簇对应功能专门化的皮层区,同一簇内节点高度连通,簇与簇之间稀疏连通。生物大脑皮层网络展示最大的结构复杂性和动力学复杂性,从而保证对丰富感知环境的响应,并编码最大数量的脑功能状态。

Kaiser研究<sup>[15,16]</sup>的出发点是对真实的生物大脑皮层网络建模。鉴于客观存在的诸如头骨边界、容积等立体空间约束,他们的模型假定一个先驱节点是一个簇的种子,先驱节点的个数对应最终生成网络中簇的个数。假定所有先驱节点在一个三维立体空间中高度对称地确定性布置。基于时间窗口和空间距离机理,归属同一个先驱节点的节点之间优先连接,

距离接近的节点之间优先连接.得到的网络具有诸如小世界、无标度分布、组件最优布置以及多簇等特征.真实地模拟了猫和短尾猿的部分大脑皮层网络.而对于工程应用的神经网络模型而言,不存在空间上的任何客观约束.我们在一个二维平面内任意布置  $n$  个先驱节点.实现的网络生长算法描述如下.

1)假定网络规模为  $N$ ,先驱节点的个数为  $n$ .在一个长和宽都为 1 的方格区域中随机布置这  $n$  个先驱节点.记录  $n$  个先驱节点的空间位置,并设定先驱节点的时间窗口<sup>1)</sup>宽度  $twwidth$ .计算每个先驱节点的时间窗口确定的概率  $p_{ij}, i = 1, \dots, n, j = 1, \dots, N$ . 2)随机产生一个新节点  $u$ .该节点与已存在节点空间位置不重叠.根据最近空间距离判断该新节点归属先驱节点  $n_u$ .若存在不止一个的先驱节点,则从中随机选择一个作为先驱节点.计算节点  $u$  由其先驱节点  $n_u$  决定的依赖时间窗口的生长概率  $p_{time}^{n_u, u}$ . 3)假设已存在旧节点个数为  $k$ .计算新节点  $u$  依赖空间距离的生长概率  $p_{dist}^{u, i} = \beta e^{-\alpha \times dis(u, i)}, i = 1, \dots, k$ .  $dis(u, i)$  是该新节点  $u$  和已存在节点  $i$  之间的欧氏距离.  $\alpha$  和  $\beta$  是两个定常距离参数. 4)计算节点  $u$  的生存概率,即与  $k$  个旧节点建立连接的概率  $p_{survive}^{u, i} = p_{dist}^{u, i} \times p_{time}^{n_u, u} \times p_{time}^{n_u, i}, i = 1, \dots, k$ . 5)  $rand$  是  $[0, 1]$  区间内的一个均匀随机数,如果  $rand < p_{survive}^{u, i}$ ,则新节点与已存在节点  $i$  连接.该节点  $u$  被接受为网络中的一员,  $k = k + 1$ .如果没有发生连接,则回到 2)重复执行.直到生成  $k = N$ . 5 元组网络生长参数  $\{N, n, twwidth, \alpha, \beta\}$  描述网络整个生长过程.

根据复杂网络理论中网络平均特征路径长度、聚集系数以及度分布概念<sup>[20]</sup>,分析本文网络生长算法生成的网络的拓扑特征.网络的平均特征路径长度  $l$  和簇集系数  $C$  分别由下式计算得到:

$$l^{-1} = \frac{1}{2} \frac{1}{N(N-1)} \sum_{i \neq j} d_{ij}^{-1}, \quad (4)$$

$$C_i = \frac{1}{\binom{k_i}{2}} E(\Gamma_i), C = \frac{1}{N} \sum_i C_i, \quad (5)$$

其中  $i, j = 1, \dots, N, d_{ij}$  是节点  $i$  和节点  $j$  之间的特征路径长度(也是最短路径长度).  $E(\Gamma_i)$  是节点  $i$  的邻居节点之间真实存在的连接个数.  $\binom{k_i}{2}$  表示节

点  $i$  的邻居节点之间可能存在的连接个数.  $C_i$  表示节点  $i$  个簇集系数.对所有  $N$  个节点的簇集系数求平均,得到网络的簇集系数  $C$ .

网络的度分布函数  $p(k)$  表示网络中具有  $k$  个连接的节点个数占网络全部节点数的百分比  $p(k)$ ,换句话说就是随机选择一个节点,该节点与网络中其他  $k$  个节点相连接的概率为  $p(k)$ .网络度分布描述网络中连接的均匀程度.网络连通密度  $\rho$  定义为网络中存在的连接个数与可能的所有连接个数的比率.

$\{N, n, twwidth, \alpha, \beta\}$  的不同取值,导致所生成的网络具有不同的拓扑特征.5 个参数依次取值  $\{100, 1, 1, 0, 1\}$ ,得到一个全连通网络( $\rho = 1, C = 1$ ).5 个参数依次取值  $\{100, 100, 0.01, 0, 1\}$ ,得到一个几乎不存在连通的稀疏网络( $\rho = 0.001, C = 0$ ).当该组参数在相应区间分别取值时,生成介于全连通网络和几乎不连通网络之间的复杂网络.图 2 和图 3 分别给出两个实现实例.

图 2 是  $N = 50, n = 5, twwidth = 0.6, \alpha = 40, \beta = 0.2$  设置下所生成的多簇复杂网络的连通图.该图使用复杂网络分析工具 Pajek<sup>[21]</sup>得到.从图中可以看出,该网络由以标号为 1, 2, 3, 4, 5 的 5 个先驱节点为核心的 5 个簇耦合构成.同簇内高度连通,簇与簇之间稀疏连通,并且存在少数长程连接.该多簇网络具有小世界属性.产生一个与该网络相同规模  $N$ 、相同连通密度  $\rho$  的完全随机网络<sup>[7]</sup>.多簇复杂网络的  $C$  值(0.32958)大于该完全随机网络的  $C$  值(0.16281),  $l$  值(2.069)与该完全随机网络的  $l$  值(1.99196)相当.  $C$  和  $l$  分别根据(4)式和(5)式计算得到.

图 3 表达  $N = 1000, n = 100, \alpha = 40, \beta = 0.2, twwidth$  分别取值 0.01, 0.2, 0.6 和 0.8 情形下,所得多簇复杂网络的度分布.图中左侧满足无标度网络具有的少数节点高度连通,大多数节点稀疏连通的无标度属性,图中右侧满足规则网络的度分布特征<sup>[22]</sup>.该度分布函数不同于文献[16]中无标度网络的度分布函数.图 3 表明,所生成多簇网络是无标度网络和规则网络的一个混合网络.图 2 和图 3 表明,本文 5 元参数组描述的网络生长算法能够自然生成一类多簇复杂网络.

1)这里使用文献[16]中时间窗口概念和时间窗口函数.

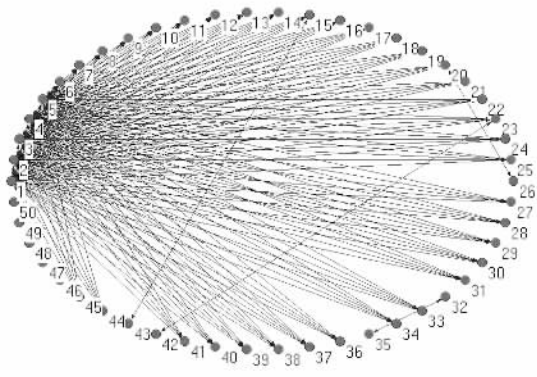


图 2 多簇复杂网络连通图

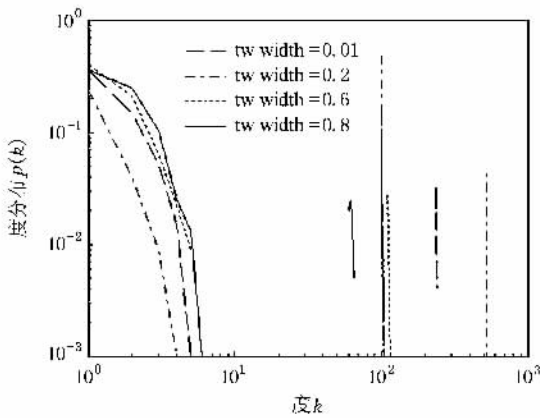


图 3 多簇复杂网络的度分布示意图

## 4. 实验设计与实现

调整网络生长参数取值,将生成的多簇复杂网络用作 ESN 的 DR. 将多簇 ESN 用于 MG 混沌时间序列预测. 目的是分析网络生长参数的不同取值,导致该多簇 ESN 的建模性能随之变化的规律,即网络生长参数与建模性能之间的关联关系.

### 4.1. 用于 MG 混沌时间序列预测

MG 混沌时间序列预测是非线性系统预测的标准问题之一. MG 动力系统如下式:

$$\frac{dx}{dt} = \frac{0.2x(t-\tau)}{1+x(t-\tau)^0} - 0.1x(t), \quad (6)$$

当  $\tau \geq 17$  时,其状态表现出非线性混沌行为<sup>[23]</sup>. 第 84 步前向预测误差,通常作为所建立黑箱模型质量

的指标.

为了便于比较,取  $\tau = 17$ ,采用文献[9]中的数据集和评价指标,即采用

$$\text{NRMSE}_{84} = \left[ \frac{\sum_{i=1}^k (d^i(+84) - y^i(+84))^2}{k\sigma^2} \right]^{1/2} \quad (7)$$

描述的 NRMSE 作为性能评价指标.

实验算法如下: 1) 生成网络. 设置多簇网络生长参数  $\{N, n, \text{twwidth}, \alpha, \beta\}$ , 调用第 3 部分的多簇网络生长算法, 将所生成网络的邻接矩阵设置为 (1) 式中的  $W$ , 并初始化  $W^{\text{in}}, W^{\text{back}}$  在  $[-1, 1]$  区间内随机均匀取值. 这样, 得到一个初始化的多簇 ESN  $\{N, W, W^{\text{in}}, W^{\text{back}}\}$ . 2) 设置库邻接矩阵谱半径  $\rho(W)$  为 0.85. 根据 (1)–(3) 式, 完成系统辨识, 建立 MG 非线性系统黑箱模型. 3) 随机生成  $k$  个 MG 测试序列, 运行所得到的模型, 根据 (7) 式计算  $\text{NRMSE}_{84}$ , 作为本次多簇 ESN 的预测误差 (本实验  $k = 30$ ) (7) 式中  $d^i(+84)$  和  $y^i(+84)$  分别是第  $i$  个测试序列的前向第 84 步真实输出值和实际输出值,  $i = 1, \dots, k$ . 4) 保持  $\{N, n, \text{twwidth}, \alpha, \beta\}$  取值不变, 重复 1)–3) 步骤 20 次, 即 Monte Carlo 仿真实验次数为 20, 取 20 次 ( $\text{NRMSE}_{84}$ ) 的均值  $\text{mean}$  和标准差  $\text{std}^2$ ,  $i = 1, \dots, 20$ , 作为该组生长参数对应多簇 ESN 的预测性能评价. 5) 针对 5 个网络生长参数, 固定其中的三个, 改变其他两个参数取值, 重复 1)–4) 步骤, 得到网络生长参数与 ESN 预测性能评价的关系表达. 实验结果如图 4 所示.

图 4(a)(b)(c) 表明, 网络生长参数  $\{N, n, \text{twwidth}, \alpha, \beta\}$  的改变能够明显改变多簇 ESN 黑箱模型的预测误差  $\text{mean}$ . 图 4(a) 是在  $\{N, \alpha, \beta\}$  分别取值  $\{300, 40, 0.5\}$  情形下,  $\text{mean}$  随参数  $\{n, \text{twwidth}\}$  变化的关系图. 该图表明在库规模和距离参数一定的前提下,  $n$  和  $\text{twwidth}$  大约成反比例线性关系, 能够保证较低的  $\text{mean}$  值. 图 4(b) 是在  $\{N, n, \text{twwidth}\}$  分别取值  $\{300, 3, 0.3\}$  情形下,  $\text{mean}$  随参数  $\{\alpha, \beta\}$  变化的关系图. 该图表明,  $\text{mean}$  对距离参数  $\beta$  的变化比对距离参数  $\alpha$  的变化更敏感. 图 4(c) 是在  $\{\text{twwidth}, \alpha, \beta\}$  分别取值  $\{0.5, 40, 0.5\}$  情形下,  $\text{mean}$  随参数  $\{N, n\}$  变化的关系图. 对于每个  $N$ , 存在一个  $n$  能够导致 ESN 具有最小的  $\text{mean}$ ; 同样对于每个  $n$ , 存在一个  $N$  能够导致 ESN 具有最小的  $\text{mean}$ .

$$2) \text{mean} = \frac{1}{20} \sum_{i=1}^{20} (\text{NRMSE}_{84})_i, \quad \text{std} = \left[ \frac{1}{20-1} \sum_{i=1}^{20} [(\text{NRMSE}_{84})_i - \text{mean}]^2 \right]^{1/2}.$$

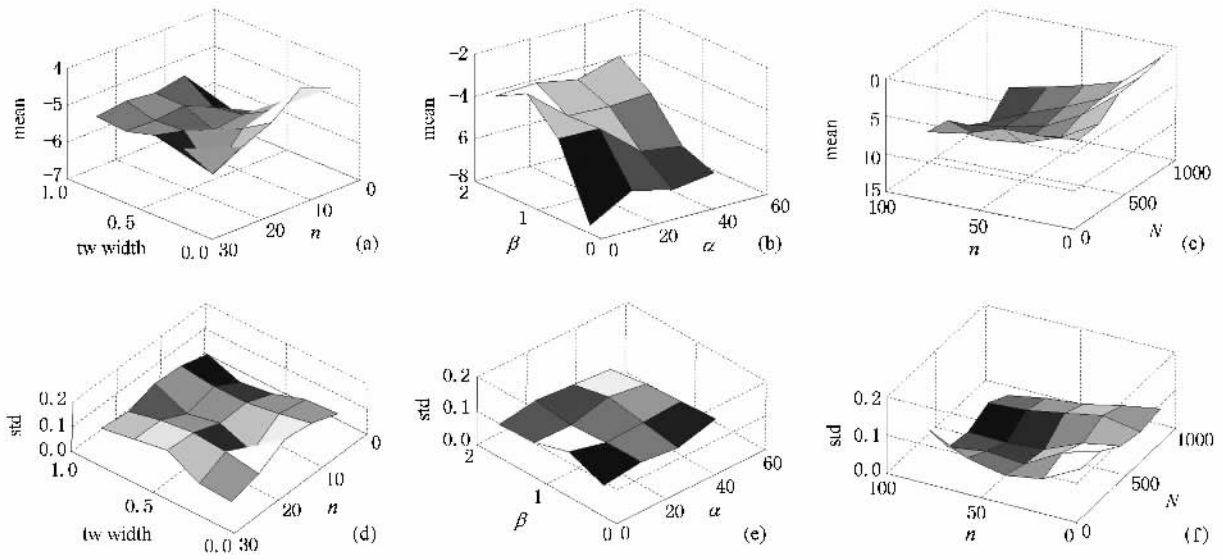


图4 网络生长参数与ESN预测性能评价的关系示意图。(a)(b)(c)是网络生长参数 $\{N, n, tw\text{width}, \alpha, \beta\}$ 与多簇ESN预测误差的mean的关系图。(d)(e)(f)是与(a)(b)(c)一一对应的网络生长参数 $\{N, n, tw\text{width}, \alpha, \beta\}$ 与多簇ESN预测误差的std的关系图。(a)(b)(c)三个子图的z坐标是对数刻度。实验在Matlab7中完成。

此外,在图4(c)中,当 $\{N, n, tw\text{width}, \alpha, \beta\}$ 分别取值 $\{500, 40, 0.5, 40, 0.5\}$ 情形下,多簇ESN预测误差mean约为0.0002,这个结果接近文献[9]中 $N$ 取1000的标准ESN的预测性能。这表明多簇ESN比Jaeger的标准ESN能更有效地处理MG混沌时间序列预测任务。

图4(d)(e)(f)表明,20次独立随机实现的多簇ESN预测误差具有小的std值,说明 $(NRMSE_{84})_i$ 平稳分布,  $i = 1, \dots, 20$ 。

同时考察每组参数对应的mean和std,发现具有较低预测误差mean的多簇ESN,其误差std也较小,这在图4(a)(d)和图4(c)(f)中表现较为明显。正如第3部分所述,这些参数值下生成的网络正是类似真实皮层网络的具有小世界、无标度和多簇等属性的多簇复杂网络,正是这样的拓扑保证多簇ESN具有较低的预测误差。而不具备这样拓扑的动态库,mean和std都比较大。对于工程应用,小方差能够保证在针对特定任务设计最佳动态库时,随机生成较少数量的动态库,就能从中获取具有最小预测误差的动态库。图4(b)(e)表达距离参数 $\alpha$ 和 $\beta$ 与ESN预测误差的关系。这两个子图表明,预测误差几乎不受参数 $\alpha$ 影响,这与Liebald的结果<sup>[24]</sup>是一致的。

该组实验表明,网络生长参数能够有效地改变动态库的拓扑,进而影响ESN对MG混沌时间序列

预测的性能。用于MG混沌时间序列预测的具有小世界、无标度和多簇属性的多簇ESN具有小误差小方差优势。

## 5. 讨 论

本质上,ESN库中个体状态实现一组基 $\{x_1(n), \dots, x_N(n)\}, n = 1, \dots, K$ ,这组基张成输入激励的高维表达空间。读取器将期望的输出投影到该组基上,投影大小即为输出权重向量 $(W^{\text{out}})_{N \times 1}$ 。ESN建模的关键问题在于寻找这样一个能够充分表达输入激励的高维状态空间。换句话说,已知期望输出,寻找一组基,该组基的线性组合能够充分逼近期望输出。现有文献常常是将ESN库中非线性个体在其当前状态的一个小邻域内线性化,从而以一个时变线性系统的组合来近似ESN高维非线性系统。从线性系统理论出发,考察这组线性系统的特征值分布,间接分析ESN高维非线性动力学特征。文献[11]中生成的ESN库的最大特征值分布服从幂率,这不仅改善库的回响状态属性,而且改善网络对随机噪声的鲁棒性。Principe等定义度量ESN动力学丰富程度的回响状态平均熵。平均状态熵最大值的情形下,线性化的ESN库具有最均匀的极点分布,这保证库中个体基函数具有最均匀取值的时间常数。这组相关度最小的基函数张成库最大状态空间,从而保证库

对输入激励的充分表达能力<sup>[25]</sup>. 这些从线性系统的角度设计和分析 ESN 高维非线性系统的方法, 较好地完成一些非线性系统的建模.

对于本文的多簇 ESN, 定义 DR 状态轨迹在  $N$  维状态空间中的线性复杂性, 即  $\Omega$  复杂性, 用它刻画 DR 的表达能力.  $\Omega$  复杂性已经被有效地用于描述空间上分布的多通道 EEG 信号序列的空间复杂度<sup>[26-28]</sup>. ESN 训练过程中得到的收集状态矩阵  $[\cdot]_{N,C}$  是对 DR 中  $N$  个节点  $C$  次采样的结果. 这  $C$  次采样得到的向量  $X$  取值构成库  $N$  维状态空间中的一条轨迹. 定义状态协方差矩阵如下:

$$C_{ov} = \frac{1}{C}([\cdot]_{N,C}[\cdot]_{N,C}^T). \quad (8)$$

计算协方差矩阵  $C_{ov}$  的特征值  $\lambda_1, \dots, \lambda_N$ , 对这些特征值进行归一化,  $\lambda'_i = \lambda_i / \sum_i \lambda_i$ , 得到  $\lambda'_1, \dots, \lambda'_N$ , 进而得到  $\Omega$  复杂性度量

$$\ln \Omega = - \sum_i \lambda'_i \ln \lambda'_i, \quad (9)$$

$\Omega$  具有明确的物理意义. 该值从总体上描述库状态轨迹的空间线性复杂性. 理想情形下,  $C_{ov}$  是一个对角矩阵, 具有  $N$  个相同的特征值, 表明  $N$  个神经元个体活性过程完全不同步, 状态互不相关, 作为  $N$  个主成分分别分散到  $N$  个主轴上, 构成 DR 状态空间的一组正交基. 此时  $\Omega$  对应其最大值, 库具有最好的输入激励表达能力. 另外一个极端情形下,  $C_{ov}$  是除一个特征值不等于零之外, 其余全等于零, 表明  $N$  个神经元活性过程完全同步, 状态高度互相关, 主成分完全集中到该特征值对应的主轴上. 此时  $\Omega$  值为零, 库具有最差的输入激励表达能力.

在计算预测误差  $\text{mean}$  的同时, 根据 (8) 式和 (9) 式度量 ESN 库状态的  $\Omega$  复杂性. 结果如图 5 所示.

结合图 4 和图 5, 可见  $\Omega$  值与  $\text{mean}$  具有明显的

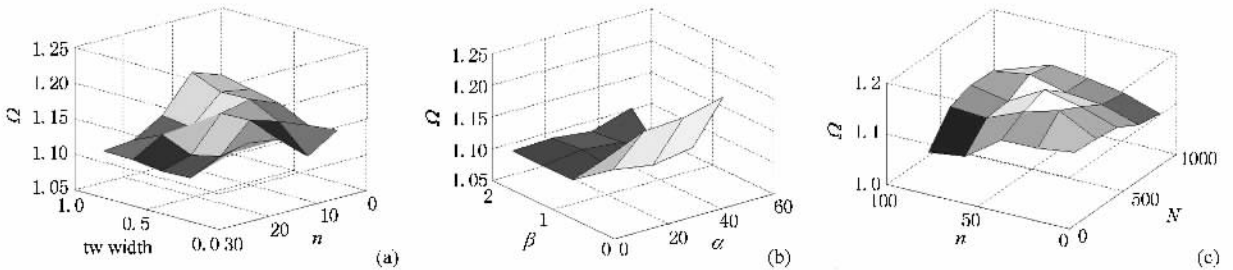


图 5 网络生长参数与 ESN 库状态轨道空间复杂度  $\Omega$  的关系图 (a)(b)(c) 与图 4(a)(b)(c) 一一对应

相关关系.  $\Omega$  值越大,  $\text{mean}$  值越小, 说明多簇 ESN 动态库表达能力越强, 其预测性能也就越好; 反之亦然. 同时可以看到, 不同的生长参数取值, 生成不同拓扑特征的复杂网络, 具有不同的表达能力. 所有  $\Omega$  都是 20 次 Monte Carlo 仿真实验结果的平均.

该组实验表明,  $\Omega$  复杂性能够有效度量 ESN 动态库对输入激励信号的表达能力. DR 的信号表达能力直接影响 ESN 模型的预测性能. 网络生长参数的不同取值导致 ESN 建模性能的改变, 本质上是通过改变 DR 的拓扑, 造成该 DR 对输入激励信号的表达能力改变而导致的结果.

## 6. 结 论

本文在 Kaiser 多簇网络生长算法基础上, 实现

了 5 个生长参数表达的多簇复杂网络. 在该组生长参数特定取值情形下, 所得网络具有小世界和无标度属性, 并且表现出多簇组织特征. 将该多簇复杂网络作为 ESN 动态库使用, 实现了对 MG 混沌时间序列的准确预测. 定义了刻画多簇 ESN 动态库信号表达能力的  $\Omega$  复杂性度量. 实验表明, 对于 MG 混沌时间序列预测任务, 多簇 ESN 比 Jaeger 的标准 ESN 更有效, 比不具有多簇复杂网络属性的 ESN 更有效. 不同取值的网络生长参数所产生的动态库具有不同的  $\Omega$  复杂性. 具有较高  $\Omega$  复杂度取值的多簇 ESN 具有更好的预测性能. 本研究为 ESN 用于混沌时间序列预测提供了一个有效的建模方法.

- [ 1 ] Jiang B , Wang H Q , Li X *et al* 2006 *Acta Phys. Sin.* **55** 3985 ( in Chinese ) [ 姜 斌、王宏强、黎 湘等 2006 物理学报 **55** 3985 ]
- [ 2 ] Ding G , Zhong S S 2007 *Acta Phys. Sin.* **56** 1224 ( in Chinese ) [ 丁 刚、钟诗胜 2007 物理学报 **56** 1224 ]
- [ 3 ] Gao C H , Zhou Z M 2004 *Acta Phys. Sin.* **53** 4092 ( in Chinese ) [ 郜传厚、周志敏 2004 物理学报 **53** 4092 ]
- [ 4 ] Haykin S , Principe J 1998 *IEEE Signal Processing Magazine* **15** 66
- [ 5 ] Ren X M , Rad A B 2007 *IEEE Trans. Neural Networks* **18** 1536
- [ 6 ] Feldkamp L , Prokhorov D , Eagen C , Yuan F 1998 *Nonlinear Modeling : Advanced Black-Box Techniques* ( Boston , MA : Kluwer Academic Publishers ) p29
- [ 7 ] Jaeger H 2001 *GMD Technical Report* 148 ( German National Research Center for Information Technology )
- [ 8 ] Maass W , Natschlager T , Markram H 2002 *Neural Computation* **14** 2531
- [ 9 ] Jaeger H , Haas H 2004 *Science* **304** 78
- [ 10 ] Han M , Shi Z W , Guo W 2007 *Acta Phys. Sin.* **56** 43 ( in Chinese ) [ 韩 敏、史志伟、郭 伟 2007 物理学报 **56** 43 ]
- [ 11 ] Deng Z , Zhang Y 2007 *IEEE Trans. Neural Networks* **18** 1364
- [ 12 ] Xue Y , Yang L , Haykin S 2007 *Neural Networks* **20** 365
- [ 13 ] Bush K , Anderson C 2006 *Neural Information Processing Systems* ( Whistler B. C. )
- [ 14 ] Jaeger H 2005 *Proceedings of IEEE IJCNN* **3** 1460
- [ 15 ] Kaiser M , Hilgetag C C 2007 *Neurocomputing* **70** 1829
- [ 16 ] Nisbach F , Kaiser M 2007 *Eur. Phys. J. B* **58** 185
- [ 17 ] Chen S , Chang S J , Yuan J H *et al* 2001 *Acta Phys. Sin.* **50** 674 ( in Chinese ) [ 陈 戎、常胜江、袁景和等 2001 物理学报 **50** 674 ]
- [ 18 ] Sporns O , Tononi G , Edelman G M 2000 *Cereb. Cortex* **10** 127
- [ 19 ] Kaiser M 2007 *Phil. Trans. R. Soc. A* **365** 3033
- [ 20 ] Newman M E J 2003 *SIAM Review* **45** 167
- [ 21 ] Batagelj V , Mrvar A 2003 *Graph Drawing Software* ( Berlin : Springer ) p77
- [ 22 ] Wang X F , Li X , Chen G R 2006 *Complex Network Theory and Application* ( Beijing : Qinghua University Press ) p18 ( in Chinese ) [ 汪小帆、李 翔、陈关荣 2006 复杂网络理论及其应用 ( 北京 清华大学出版社 ) 第 18 页 ]
- [ 23 ] Tan W , Wang Y N , Zhou S W *et al* 2003 *Acta Phys. Sin.* **52** 795 ( in Chinese ) [ 谭 文、王耀南、周少武等 2003 物理学报 **52** 795 ]
- [ 24 ] Liebal B 2004 *Exploration of effects of different network topologies on the ESN signal cross-correlation matrix spectrum* ( Bachelor of Science thesis , International University Bremen )
- [ 25 ] Ozturk M C , Xu D M , Principe J C 2007 *Neural Computation* **19** 111
- [ 26 ] Wackermann J 1999 *International Journal of Psychophysiology* **34** 65
- [ 27 ] Liu J P , Zheng C X 2005 *Journal of Biomedical Engineering* **22** 906 ( in Chinese ) [ 刘建平、郑崇勋 2005 生物医学工程学杂志 **22** 906 ]
- [ 28 ] Kondakor I , Brandeis D , Wackermann J *et al* 1997 *Neuroscience Letters* **226** 49

# Multiple clusters echo state network for chaotic time series prediction<sup>\*</sup>

Song Qing-Song<sup>†</sup> Feng Zu-Ren Li Ren-Hou

( Systems Engineering Institute , Xi 'an Jiaotong University , Xi 'an 710049 , China )

( Received 7 January 2008 ; revised manuscript received 9 December 2008 )

## Abstract

The chaotic time series prediction problem is considered. A novel type of cortex-like neural network model , i. e. multi-clusters echo state network model( MCESN ) , regulated by a group of five growth-factors , is proposed. It is shown that characters of MCESN ' topology can be effectively determined by the growth-factors group ; and that it is the MCESN possessing both small-world and scale-free properties of complex network that corresponds to the better prediction performance. In addition , Monte Carlo simulation experiments show that MCESN not only can be trained by easy algorithm , but also can achieve higher accuracy and less standard deviation prediction results than classical echo state networks.

**Keywords** : chaotic time series prediction , echo state networks , complex networks , Omega complexity

**PACC** : 8710 , 0547

---

<sup>\*</sup> Project supported by the Specialized Research Fund for the Doctoral Program of Higher Education of China ( Grant No. 20050698032 ) , the State Key Development Program for Basic Research of China ( Grant No. 2007CB311006 ) , and the National High Technology Research and Development Program of China ( Grant No. 2006AA04Z222 ) and the National Natural Science Foundation of China ( Grant No. 60875043 ).

<sup>†</sup> E-mail : qssong@sei.xjtu.edu.cn