

# 基于优化回声状态网络的股价预测研究

张宁致, 周佳丽, 孙武军

(南京大学商学院, 江苏 南京 210093)

**摘要:** 股票市场是一个充斥着各种噪声的动态非线性系统, 能够精确地对其进行预测是一项具有挑战性的任务。本文构建的 BCC-ESN 模型, 是运用细菌群体趋药性算法 (BCC) 来优化回声状态网络 (ESN) 的权值结构, 在继承 ESN 优良性质的同时, 具有更高的模型预测能力。实验证明, BCC-ESN 模型比前馈神经网络具有更好的学习和预测能力。经对上证指数进行短期价格预测, 结果与 BP 网络、Elman 网络和 ESN 网络进行比较, BCC-ESN 模型精度明显优于其他三种网络预测。同时, 在运算效率上, BCC-ESN 模型继承了 ESN 的运算优势, 明显优于其他神经网络预测模型, 是一种切实可行、高效的预测算法, 尤其在股票时间序列预测中有广泛的实用价值。针对 BCC-ESN 模型在训练预测中遇到的问题, 比如耗费时间过长和过度拟合问题, 本文亦提供了简单易行的思路和方法。

**关键词:** 回声状态网络; 细菌群体趋药性算法; 股票价格预测; 算法优化

**中图分类号:** F830.91 **文献标识码:** A **文章编号:** 1004-6062(2014)01-0094-08

## 0 引言

股票市场是一个充斥着各种噪声的动态非线性系统, 容易受到各种因素 (包括政治事件、宏观经济数据、投资者情绪等) 的影响。针对股票市场表现出的不同特点, 人们提出了多种多样的分析和预测方法, 比如基本面分析法、技术分析法和时间序列分析法等。然而, 高噪声、严重非线性和投资者的盲目任意性等决定了股票价格预测的复杂度和难度, 这也就制约了传统的计算方法在股票市场上的运用。

计算智能是受到大自然智慧和人类智慧的启发而设计出来的一类算法的统称。这些算法将生物智慧、自然界的规律通过计算机程序实现, 得到最优化算法, 包括人工神经网络、进化计算、群体智能、人工免疫系统、模糊系统、粒子群优化算法、分布估计算法、Memetic 算法、模拟退火算法和禁忌搜索算法等。计算智能为股票市场的预测提供了诸多方法, 这些方法在复杂和变化的环境中能够或易于实现自适应机制的研究, 而且在处理海量数据模拟方面拥有大量成功的案例。但不可否认的是, 单个智能算法无论在理论上还是实际运用当中都存在一定的局限性, 例如在预测精度、运行效率

等方面。因此, 算法之间的优化组合, 如运用遗传算法对神经网络的权值进行优化, 将细菌趋药性优化算法加入到神经网络的训练过程中, 这在一定程度上都可以提高算法的效率。当前的智能算法发展趋势正是发展各种算法之间的混合方法, 因为没有任何一种典型方法能够在所有场合都优于其他方法。

本文秉承了这一思想, 以一种特殊的神经网络——回声状态网络为基础, 使用细菌群体趋药性算法对其权值结构进行了优化, 并利用这一优化模型对上证指数进行了训练和预测。结果同 BP 网络、Elman 网络进行比较, 无论是预测精度还是运行效率, 都有了大幅度的改善。

## 1 文献综述

各种智能算法在证券市场上的运用, 无论是对证券市场方向的预测、价格变化的预测以及提供投资决策, 都取得了不错的效果。表 1 列举了部分算法在证券市场上的应用。需要注意的是, 由于预测区间的不同、研究市场的不同以及构建误差的差别, 表格中的误差结果并没有可比性。

表 1 计算智能在证券市场上的运用

作者	方法	研究市场	样本空间	误差
Wang YF <sup>[1]</sup>	模糊灰色系统 (FGS)	台湾证券市场	2000-09 至 2001-04	平均偏离率 < 9%
Shen W 等人 <sup>[2]</sup>	人工鱼群算法 (AFSA)	上海证券市场	2006-01 至 2006-03	平均误差 = 0.5395%
Majhi R 等人 <sup>[3]</sup>	细菌粒子群优化算法 (BFO)	标准普尔 500	1994-01 至 2006-10	平均绝对误差率 = 0.6623
Chun SH 等人 <sup>[4]</sup>	动态自适应组合推断 (DA-CBR)	韩国证券市场	2000-01 至 2004-06	平均绝对误差率 = 1.5914
Tan TZ 等人 <sup>[5]</sup>	遗传互补学习算法 (GCL)	新加坡证券市场	1995-09 至 2004-12	均方根误差 = 0.0032

收稿日期: 2012-02-04 修回日期: 2012-12-20

基金项目: 国家自然科学基金资助项目 (70972032); 国家社科重点资助项目 (08AJY029); 江苏高校优势学科建设工程项目资助项目 (PAPD)

作者简介: 张宁致 (1988—), 男, 南京大学商学院金融与保险学系研究生, 研究方向为金融工程与风险管理。

在众多智能算法当中,神经网络算法颇受关注。它是一种非常成熟的用于解决非线性问题的计算智能方法,结构简单从而容易理解掌握,而且它可以处理大量包含各种噪声的数据,而不像金融时间序列模型,需要数据服从诸如正态分布和稳定性这样的假定。神经网络不仅能够学习训练集的例子,且能从训练集中提炼出某种一般性原理、规律,具有很强的非线性函数拟合特性,这对于预测短周期内股指波动有较强的适用性。姚洪兴和盛昭瀚运用小波神经网络并结合相空间重构方法对股指进行了预测,得到较好的结果<sup>[6]</sup>。朱林利用粗集、神经网络和遗传算法相结合得到的混合杂交模型对股指进行预测,虽然他们使用了一定的算法优化网络结构或原始数据,从而使得预测误差得以减小,可是传统神经网络的缺点,比如收敛速度慢、存在局部最小点等,并没有得到有效地克服<sup>[7]</sup>。苏治等人运用神经网络中的 MLP 网络、JCP 网络以及部分回馈 Elman 网络对证券价格非线性动态特征及可预测性进行研究,认为相比于 RW 模型、线性 AR 模型和 STAR 模型,基于神经网络的模型的“盯市”投资策略能获得比“买入持有”投资策略更高的平均净收益<sup>[8]</sup>。王志刚等人选择了基于移动平均规则的人工神经网络模型对中国股市进行了预测,精度有所提高<sup>[9]</sup>。

回声状态网络(ESN)是一种特殊的神经网络,相比其他传统的神经网络,具有很好的短期记忆能力<sup>[10]</sup>。其简单的训练过程及独特的隐层构造方式使得训练效率大幅提高,从而在非线性和混沌时间序列预测、随机序列建模、图形模式识别等方面得到广泛应用,尤其在时间序列预测方面。ESN 在稳定性、全局最优性、局部最小问题和训练过程复杂性相对传统神经网络都有很大改进。并且,ESN 有着良好的函数逼近能力,可以准确的逼近非线性动态系统,它的多输入多输出处理能力使其可以很好地用于多元变量混沌时间序列模型描述。Jaeger 将其与前向型神经网络和传统的递归神经网络对比,给出了 ESN 的算法规则和训练模式<sup>[11,12]</sup>。Ishii 等人运用进化算法对 ESN 中的变量进行设定,同时认为,ESN 的优异性表现在不需要像传统算法那样需要依靠大量的经验<sup>[13]</sup>。Skowronski 和 Harri 将 ESN 网络运用在自动语言识别的分类上,证实了 ESN 对于噪声数据处理的优异特性,并讨论了储备池的权值矩阵的稀疏程度对处理噪声数据的影响<sup>[14]</sup>。Lin 运用 Hurst 指数来设定 ESN 中权值的初始值,他们同样认为初始值的选择在以往的网络设定中完全基于经验或者随机设定,这使得预测结果变得不稳定<sup>[15]</sup>。国内对于回声状态网络的运用主要集中在模型识别以及时间序列的预测上。韩敏和穆大芸提出了一种多核 Bayesian 状态回声机(MrBESN),其对多变量混沌时间序列的预测效果比传统的回声状态网络及支持向量回声机好。同时,采用 Bayesian 方法对多储备池 ESN 的输出权值及参数进行求取,解决了 ESN 网络在训练样本小于储备池维数时的过度拟合问题,改善了预测效果<sup>[16]</sup>。

另一个在本文中涉及到的算法是细菌群体趋药性算法。传统的细菌趋药性(Bacterial Chemotaxis, BC)的函数优化方

法是一种从生物行为中获得灵感的优化方法。这种优化方法旨在利用细菌在化学引诱剂环境中的运动行为来进行函数优化。BC 算法总体性能上可以与一些基于群体智能的优化算法进行比较,不过效果要差。Muller 和 Marchetto 等人优化了这一算法,并系统地描述了在 2 维和 n 维中的应用<sup>[17]</sup>。李威武等人提出了一种从细菌趋药性算法发展起来的新的基于群体智能的优化方法:细菌群体趋药性(Bacterial Colony Chemotaxis, BCC)算法。假定细菌在引诱剂环境中通过一些简单的假设方式进行通信。细菌不仅使用自己的运动位置信息也利用其他细菌的位置信息进行函数优化<sup>[18]</sup>。Zhang 和 Wu 应用这一优化算法与 BP 网络结合,应用在对标准普尔 500 指数的训练和预测,并与传统的 BP 网络进行对比, BCC 算法有效地减少了运算时间和提高了预测精度<sup>[19]</sup>。

综上所述,传统的神经网络算法虽然在预测股票价格上体现出了优势,但其收敛速度慢、存在局部最小点等缺点并没有得到有效地克服。而算法之间的优化组合却能有效解决这一问题。本文提出了一种优化 ESN 连接权值的方法——BCC-ESN 模型,该模型生成 ESN 连接权值的初始种群,然后利用 BCC 算法对其进行优化筛选,以确定满足问题条件下最优的权值结构,这在一定程度上大幅提高了网络的预测精度。同时,就同一系列类似问题进行训练和预测时, BCC-ESN 模型可以保存 BCC 筛选后的权值结构,就数据样本进行训练和预测,也就保证了 BCC-ESN 模型的效率。本文结构如下:首先介绍回声状态网络(ESN)和细菌群体趋药性算法(BCC);然后将 BCC 算法运用在 ESN 权值优化上,构建 BCC-ESN 模型;最后将这个模型运用在上证指数的训练和预测上,并与传统的 BP 网络和 Elman 网络进行比较。

## 2 BCC-ESN 模型构建

### 2.1 回声状态网络(ESN)

标准的 ESN 由输入层、中间层和输出层构成,其中  $K$  个输入节点与中间层  $N$  个节点连接,中间层  $N$  个节点与输出层  $L$  个节点连接,同时输出层  $L$  个节点与中间层  $N$  个节点反向连接。与一般的神经网络相比,ESN 有大量稀疏分布在中间层的节点。此外,在 ESN 网络中还允许输入层与输出层之间的连接和输出层自身的连接。如图 1 所示。

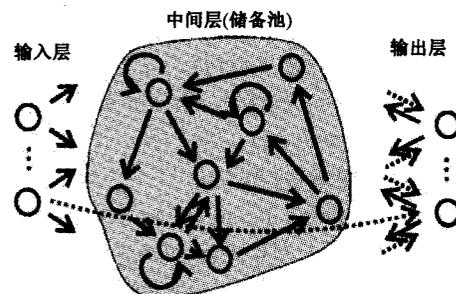


图1 ESN的基本构造

ESN 的核心结构是一个随机生成的、且保持不变的储备池(中间层)。储备池使用大规模随机连接的递归网络,取代

经典神经网络中的隐层,从而简化网络的训练过程。这个储备池一般设定为一个稀疏矩阵,其输出权值是唯一需要调整的部分。这个储备池是随机生成的、大规模的、稀疏连接的递归结构,一般要比传统的递归神经网络或者前馈网络的隐层规模大得多,大约为几十、几百左右。内部连接权值的稀疏度(S.D)一般取值为2%~5%,这是保证储备池丰富动态特性的特征之一。

对于  $n$  时刻,输入单元的序列为  $u(n) = [u_1(n), u_2(n), \dots, u_k(n)]$ , 中间层的状态序列为  $x(n) = [x_1(n), x_2(n), \dots, x_N(n)]$ , 输出单元的序列为  $y(n) = [y_1(n), y_2(n), \dots, y_L(n)]$ 。

中间层状态  $x(n+1)$  的更新方程为:

$$x(n+1) = f(W^{in}u(n+1) + Wx(n) + W^{back}y(n)) \quad (1)$$

其中  $W^{in}$  指的是输入层到中间层的权值向量,而  $W$  代表的是中间层之间构成的权值向量,而  $W^{out}$  代表的是从输出层返回到中间层的权值向量。而 ESN 的输出由下式得到:

$$y(n+1) = f^{out}(W^{out}(u(n+1), x(n), y(n))) \quad (2)$$

ESN 的一个重要特性就是只有中间层至输出层的权值向量  $W^{out}$  需要通过学习和训练来修正,而中间层和其他权值向量的拓扑结构保持不变。这样虽然大幅缩减了训练时间,但是网络会因为权值向量的初始设定的随机性而使网络输出不稳定。而本文构建的 BCC-ESN 模型会有效解决这一问题。

ESN 作为一种特殊的递归神经网络,主要表现在具有回声状态这一特性上。回声状态是指,在特定条件下,ESN 神经网络中间层的激活状态  $x(n)$  可以表示成为一系列输入向量(这个输入向量可以是无限的)的函数。而要让网络拥有这个回声状态,需要中间层权值向量满足一定的条件。传统的 BP 网络、Elman 网络,对于各层权值的设定没有要求,一般步骤是先随机选取,然后根据训练输入进行优化,而 ESN 为了满足之前提及的回声状态,需要内部权值矩阵的谱半径小于 1,也就是要求内部权值矩阵  $W$ ,其最大特征值必须满足:

$$|\lambda_{max}| < 1 \quad (3)$$

为了满足(3)式,使得递归神经网络拥有回声状态,对于一个随机生成的内部权值矩阵,求出其最大特征值(不考虑符号):即求出  $|\lambda_{max}|$ ,更新权值向量:

$$W_{new} = \frac{1}{|\lambda_{max}|} W_{old} \quad (4)$$

这样,无论输入输出矩阵取何值,按上述方法得到的动态记忆库就是一个回声状态网络。矩阵  $W$  的谱半径小于 1,保证了递归神经网络的稳定性。

## 2.2 细菌群体趋药性算法(BCC)

细菌群体趋药性算法(BCC)是一种基于细菌趋药性算法上的优化。因为现实世界中,没有生物是完全独立地生活的。尽管细菌是很细微的生物,它们也必定通过某种方式相互联系,很多研究表明某些细菌群体在觅食过程中也有群聚现象,细菌之间通过各种方式交换食物信息。使用同伴提供的信息,细菌将能够大大扩展他们对于环境的了解从而能增

加存活的几率。基于上述思想和细菌趋药性算法(BC)结合,得到了细菌群体趋药性算法(BCC)。在 BCC 算法中,对不同的评价函数进行的试验表明,细菌趋药性优化算法不仅极大地提高了细菌趋药性算法的优化能力,而且其优化速度和精度也超过了其他一些常见的群体优化算法。

在 BCC 算法中,细菌群体间遵照以下的方式进行相互联系:菌类聚群的方式是通过释放对其它细菌起引诱剂作用的化学物质来进行联系,可以假定每一个细菌都有一定的感知范围,在这个范围内细菌可以感知其它细菌以及它们的状态。也就是说,在每次移动到新位置之前,细菌都要感知它周围的环境,试探旁边是否有其他位置更好的细菌。如果有,那么它有可能趋向移动到这些拥有较好位置细菌的中心点。

## 2.3 BCC-ESN 模型构造和算法步骤

由于 ESN 的输入权值、中间层权值和反馈权值一般取随机值或根据经验设定,而且在 ESN 网络的工作期间并不像传统的神经网络,进行梯度修正,因此网络结构存在一定的随机性,通过网络训练预测出的值也不一定能达到理想的状态。就权值向量的设定问题,Ishii 等人运用进化算法对 ESN 中的变量进行设定<sup>[13]</sup>,而 Lin 运用 Hurst 指数来设定 ESN 网络的初始值<sup>[15]</sup>。本文将 BCC 算法应用于权值的优化上,大幅提高了 ESN 网络的精确度和稳定性。图 2 展示了 BCC-ESN 模型的结构。

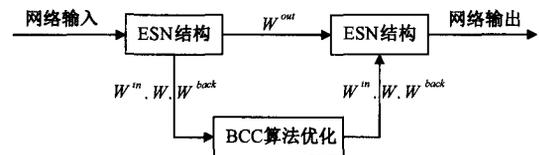


图 2 BCC-ESN 的结构图

BCC-ESN 模型的具体算法如下:

第一步:输入样本序列,将数据归一化到  $[0, 1]$  区间上,将样本序列划分为训练数据或测试数据;

第二步:设定 ESN 参数,选择储备池规模、内部权值稀疏度和内部连接权值  $W$  的谱半径。随机产生一个稀疏矩阵  $W$  作为动态记忆库,同时随机产生  $W^{in}$ 、 $W^{back}$ ,使用 ESN 训练  $W^{out}$ ;

第三步:使用 BCC 训练  $W^{in}$ 、 $W$ 、 $W^{back}$ ,使用评价指标(RMSE)计算训练误差,决定其最优值;

第四步:返回最优的  $W^{in}$ 、 $W$ 、 $W^{back}$ ,训练 ESN,然后利用其进行预测。

## 3 仿真实验

### 3.1 数据准备和网络结构

股票价格受到众多因素的影响,主要包括前几日的价格,中短期价格趋势,投资者的心理等。本文运用前 5 日的收盘价、中短期均价以及反映市场形态和投资者心理的技术指标对价格进行预测。许多学者试图对收益率和一些技术指标(比如均价)进行预测,可 Kaboudan 认为收益率和技术分析指标符合随机游走过程,相比之下,股票价格有更高的

准确率<sup>[20]</sup>。所以,本文仅对股票价格进行预测。

本文分别对小样本和大样本条件进行训练和预测,以分析在不同情形的预测精度和效率。小样本条件选取的是上证指数2007年2月至2011年3月共1000天的数据,其中训练样本和测试样本各占50%(各500天);大样本条件选取的是上证指数2000年11月至2011年3月共2500天的数据,其中训练样本占80%(2000天),测试样本各占20%(500天)。

为了验证BCC-ESN模型的预测精度和效率,这里将静态的BP网络、动态的Elman网络、和ESN与之对比。在小样本条件下,样本预测的是1日后的价格和15日后的价格,BP网络的结构为{11-50-1},Elman网络的结构同样为{11-50-1},训练迭代最大次数设定为500次,误差训练方法选择的为traingdx。ESN和BCC-ESN的储备池(中间层)节点数设为60。四种网络的输入向量相同,包括前5日每日的收盘价,和6个技术指标。在大样本条件下,样本预测的是15日后的价格。其中BP和Elman的网络结构不变,考虑到当样本数目变动时,储备池(中间层)节点数对ESN训练和预测影响较大,而对BP网络和Elman网络几乎没有影响,在这里,ESN和BCC-ESN的储备池(中间层)节点数设为200。

### 3.2 评价指标

为了分析和对比模型的实验结果,本文使用均方误差RMSE作为性能指标,对预测结果进行比较,定量表明各个模型的预测效果。设 $T_i$ 为预测变量的实际观测值, $P_i$ 为该变量的网络预测值, $N$ 为训练的样本数,则均方误差定义为:

$$RMSE = \sqrt{\frac{(P_i - T_i)^2}{N}} \quad (5)$$

均方误差反映了预测值相对于真实值的偏离程度,取值大于或等于零,当均方误差值越小时表明预测效果越好。

### 3.3 储备池谱半径的选择

上文对ESN的储备池做了介绍,除了需要满足ESN的回声状态特性之外,还需要根据不同的实际问题对储备池的谱半径进行适当调整。在谱半径小于1的前提下,谱半径具体大小对网络结构的训练有很大的影响,因为谱半径的大小关系着储备池的动态特征。一般认为,小的谱半径意味着储备池的动态特征比较明显。图3和图4展示了基于相同的样本,ESN网络在不同谱半径的设定下对于股票价格训练和预测的误差大小。通过图示可以看出,当谱半径小于0.8时,ESN网络保持较低的误差(RMSE);当谱半径位于0.8和0.9之间时,ESN网络误差以一个较小的斜率开始上升;当谱半径大于0.9时,ESN网络误差急剧上升。这说明,股票价格的动态特征比较明显,所以小的谱半径可以很好地训练股票价格。已有相关文献亦表明,当用来预测股票价格时,设定谱半径在0.1~0.5之间时可以得到较好的效果。本文运用BCC-ESN对样本进行大量实验后发现,当谱半径在0.3附

近时,网络的训练误差和测试误差都比较小,所以将BCC-ESN结构中储备池的谱半径设置为0.3。

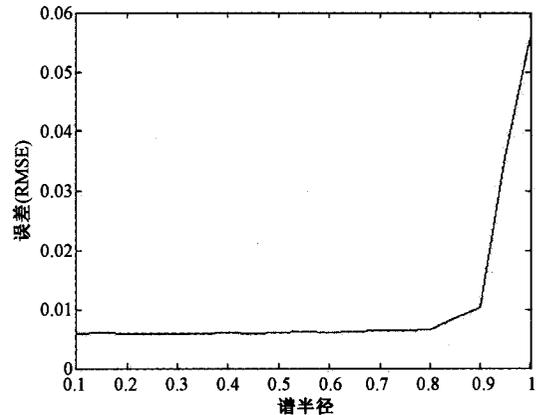


图3 BCC-ESN模型在不同谱半径条件下的训练误差

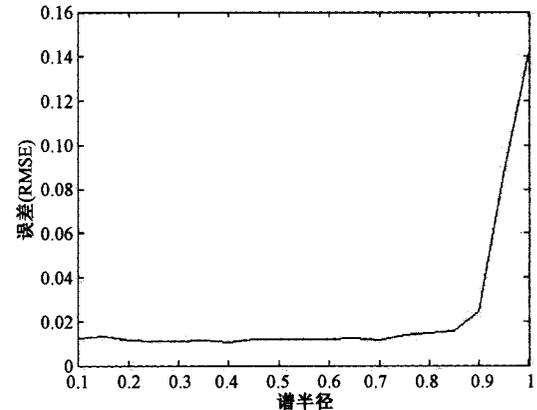


图4 BCC-ESN模型在不同谱半径条件下的预测误差

### 3.4 实验结果

仿真实验结果见图5至图12和表2至表4。由于ESN在训练预测中需要各删去100个初始瞬值,所以图中显示的ESN和BCC-ESN的训练和预测样本均减少100个值。同时,因为1日训练预测的误差图示不明显,故只给出15日训练预测图示。

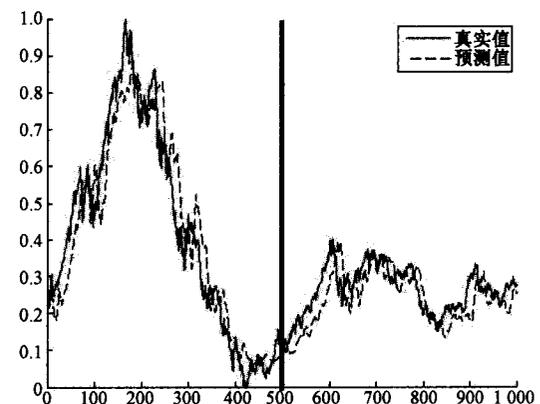


图5 BP网络对15日后的训练和预测  
(前500日训练,后500日预测)

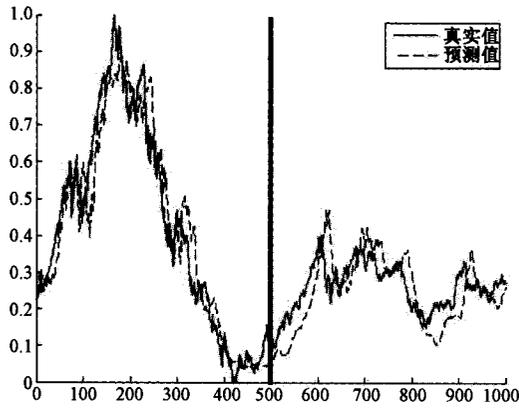


图6 Elman网络对15日后的训练和预测  
(前500日训练,后500日预测)

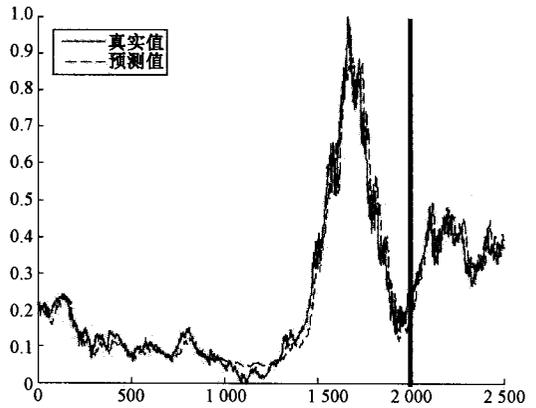


图9 BP网络对15日后的训练和预测  
(前2000日训练,后500日预测)

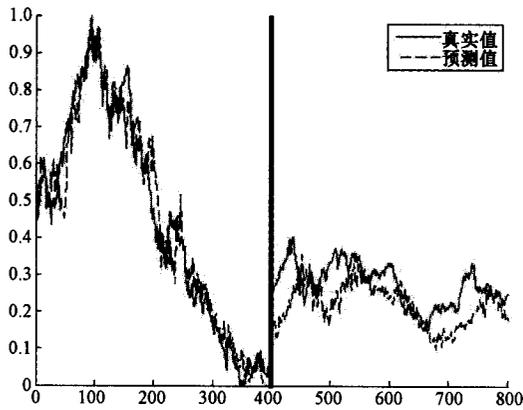


图7 ESN对15日后的训练和预测  
(前400日训练,后400日预测)

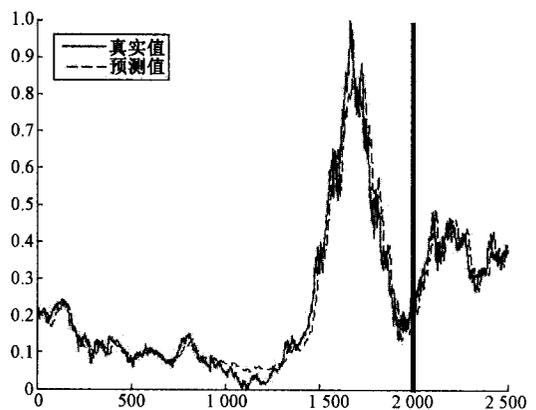


图10 Elman网络对15日后的训练和预测  
(前2000日训练,后500日预测)

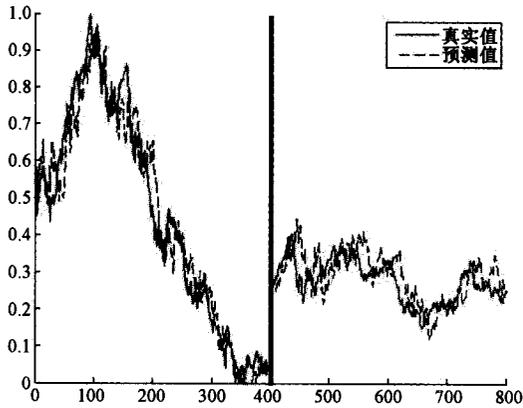


图8 BCC-ESN对15日后的训练和预测  
(前400日训练,后400日预测)

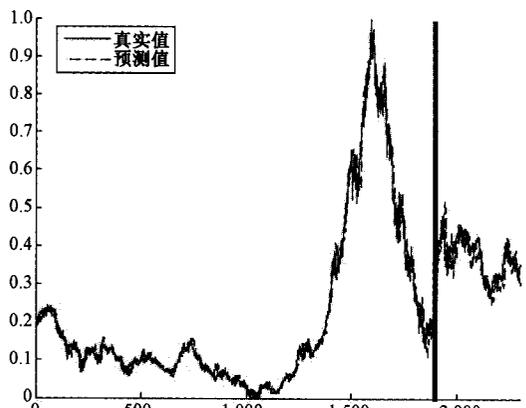


图11 ESN对15日后的训练和预测  
(前1900日训练,后400日预测)

表2 各种网络对于1日价格的训练预测误差比较  
(样本为1000日)

所用网络	训练误差 (RMSE)	测试误差 (RMSE)
BP	0.030477	0.033522
Elman	0.053679	0.070723
ESN	0.033714	0.016987
BCC-ESN	0.019616	0.010681

表3 各种网络对于15日价格的训练预测误差比较  
(样本为1000日)

所用网络	训练误差 (RMSE)	测试误差 (RMSE)
BP	0.076914	0.068683
Elman	0.086082	0.052334
ESN	0.063402	0.078413
BCC-ESN	0.059697	0.051170

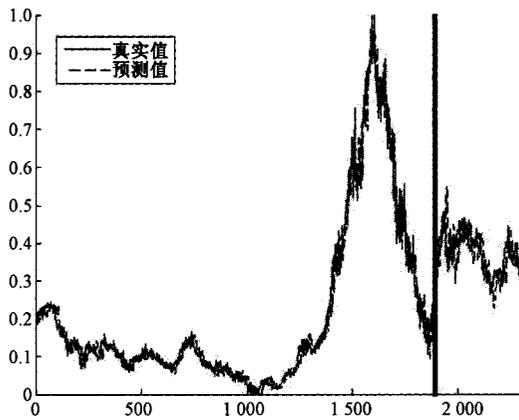


图 12 BCC-ESN 对 15 日后的训练和预测  
(前 1900 日训练, 后 400 日预测)

表 4 各种网络对于 15 日价格的训练预测误差比较  
(样本为 2500 日)

所用网络	训练误差 (RMSE)	测试误差 (RMSE)
BP	0.043584	0.040112
Elman	0.043438	0.039431
ESN	0.032779	0.044702
BCC-ESN	0.013057	0.029241

通过图表,可以看出,在 1 日价格预测的试验中,无论是训练误差还是测试误差,表现得最好的是本文构建的 BCC-ESN 模型,在训练阶段表现得最差的是 Elman 网络,在测试阶段表现得最差的是 BP 网络。BCC-ESN 良好的训练预测效果同样体现在对 15 日价格预测的试验(大样本和小样本),而此时 ESN 在精度上却失去了优势。无论是大样本还是小样本,ESN 的预测误差均是最大的,虽然预测结果存在随机性,但是原本在 1 日价格预测试验中的优势已经不再显现出来。利用 BCC 优化后的 BCC-ESN 模型很好地解决了这个问题。

表 5 样本为 1000 日数据下的耗费时间

所用网络	耗费时间(秒)
BP	10.863664
Elman	28.490031
ESN	1.641803
BCC-ESN(第一次)	230.54256
BCC-ESN(权值设定后)	1.682048

表 6 样本为 2500 日数据下的耗费时间

所用网络	耗费时间(秒)
BP	22.593612
Elman	34.826364
ESN	2.141276
BCC-ESN(第一次)	635.050007
BCC-ESN(权值设定后)	2.226372

从耗费时间上看, BP 网络比 Elman 网络耗费时间要少,

这是因为 Elman 增加了一个递归函数的运算量。相对于 BP 网络和 Elman 网络, ESN 所耗费的时间非常少,这是由于传统的神经网络(像 BP 网络和 Elman 网络),在学习的时候需要通过信息反馈来调整权值,而 ESN 的学习思路是通过学习数据的采集使储备池记忆相应的信息,最后通过线性回归过程来最小化平均误差平方值,这种训练方式属于凸优化,不必计算梯度信息,迭代次数既不是根据误差达到的精度或者设定的最大值,而是根据样本的数目。训练上的优势使得 ESN 在大样本条件和小样本条件下运行的效率均优于传统的神经网络。

可是当利用 BCC 优化 ESN 时,本文构建的 BCC-ESN 模型耗费时间出现了大幅上升,这是由于 ESN 储备池的  $num$  个节点都是自连接的,所以产生了一个  $num \times num$  的权值矩阵,其中有  $num \times num \times S.D$  个权值需要优化,权值的优化需要运算出 ESN 模型的输出,所以当存在大量数据需要优化时,模型所耗费的时间也随之升高。而当第一次运行 BCC 后,如果把权值固定下来,再利用这一优化后的 ESN 训练和预测,所耗费的时间又与本来的 ESN 相当,恢复了 ESN 训练所带来的效率提升。这就意味着如果利用 BCC-ESN 模型处理一系列同质问题时,第一次可以利用 BCC 优化 ESN 结构中的各层权值,然后保存处理问题最为精确的权值,待处理后续问题时使用。这样,既可以提高预测的精度,又能够在一定程度上克服 BCC 优化带来的耗费时间过长的问题,充分发挥 ESN 在处理问题效率上的优势。

泛化能力是神经网络学习中非常重要的方面。当 BCC-ESN 结构太过庞大时,训练过程可能会面临过度拟合问题,使得拟合优度超过了理论的最优拟合度,过度拟合的模型实际上包含了错误的信息,预测能力很差。在具体的实验中体现为,当选取较大的中间层节点时,可以在训练过程中可以达到非常精确地拟合效果,但在测试阶段,由于训练模式的过度拟合导致差的泛化误差。随着中间层节点数的增加,训练误差迅速减小,而测试误差也逐渐增大, BCC-ESN 模型的泛化能力衰减。对解决过度拟合问题,史志伟等人使用岭回归学习算法,采用交叉检验方法解决<sup>[21]</sup>,而韩敏等人使用贝叶斯线性回归解决了这一问题<sup>[16]</sup>。本文构建的 BCC-ESN 模型虽有可能面临这样的过度拟合问题,但只要将中间层的节点数目控制在合理的范围之内就可以有效解决这一问题,而且不影响模型本身的预测精度。

#### 4 结束语

人工智能已经越来越多地运用在复杂经济系统的预测上,特别是资本市场上。在新的算法不断涌现的同时,已经检验的各种智能算法正在被理论界和实务界广泛应用,特别是算法间的优化组合。一个好的、优化后的智能算法不但能够弥补单个算法的缺陷,而且在提高预测精度和效率上功效显著。本文正是用细菌群体趋药性算法(BCC)来优化回声状态网络(ESN)的权值结构,从而构建了 BCC-ESN 模型。BCC-ESN 模型,保留了 ESN 的高精度和短期记忆性的特性,

并在 ESN 基础上对 ESN 各层权值进行了优化,使得 ESN 的精度大幅提高。将其运用在上证指数的价格预测上,BCC-ESN 的预测精度明显优于 BP 等神经网络。

然而在优化过程中所费时间过长是 BCC-ESN 模型的一个问题,针对此问题本文亦提供了可行的解决思路,即在处理大量类似问题时,可先用 BCC 算法优化 ESN 权值结构,然后利用这一优化后的 ESN 进行训练和预测,这样既能够有效保证精度,更提高了效率。对于模型速度上的优化,即如何快速找到网络适应的最优点,或者如何采取一个简单的任务以完成全局的优化,是今后研究的一个方向。

### 参 考 文 献

[ 1 ] Wang YF. Predicting stock price using fuzzy grey prediction system[J]. Expert Systems with Applications, 2002, 33 ~ 39.

[ 2 ] Shen W, Guo X, Wu C, et al. Forecasting stock indices using radial basis function neural networks optimized by artificial fish swarm algorithm[J]. Knowledge-Based Systems, 2011, 378 ~ 385.

[ 3 ] Majhi R, Panda G, Sahoo G, et al. Stock market prediction of S&P 500 and DJIA using bacterial foraging optimization technique [J]. In 2007 IEEE congress on evolutionary computation, 2007, 2569 ~ 2579.

[ 4 ] Chun SH, Park YJ. Dynamic adaptive ensemble case-based reasoning: application to stock market prediction [J]. Expert Systems with Applications, 2005, 435 ~ 443.

[ 5 ] Tan TZ, Quek C, Ng GS. Brain inspired genetic complimentary learning for stock market prediction [J]. In IEEE congress on evolutionary computation, 2005(3): 2653 ~ 2660.

[ 6 ] 姚洪兴,盛昭瀚. 股市预测中的小波神经网络方法的研究 [J]. 管理工程学报,2002,16(2):32 ~ 37.

[ 7 ] 朱林,何建敏,常松. 粗集与神经网络相结合的股票价格预测模型[J]. 中国管理科学,2002,10(8):7 ~ 12.

[ 8 ] 苏治,方明,李志刚. STAR 与 ANN 模型:证券价格非线性动态特征及可预测性研究[J]. 中国管理科学,2008,16(5):9 ~ 16.

[ 9 ] 王志刚,曾勇,李平. 中国股票市场技术分析非线性预测能力的实证检验[J]. 管理工程学报,2009,23(1):149 ~ 153.

[ 10 ] Jaeger H. The "echo state" approach to analyzing and training recurrent neural networks [ R ]. Berlin: German National Research Center for Information Technology, 2001.

[ 11 ] Jaeger H. Tutorial on training recurrent neural networks covering BPPT, RTRL, EKF and the echo state network approach [ R ]. Berlin: German National Research Center for Information Technology, 2002a.

[ 12 ] Jaeger H. Short term memory in echo state networks[ R ]. GMD-Report 152, GMDGerman National Research Institute for Computer Science, 2002b.

[ 13 ] Ishii K, Zant T Van Der, Becanovic V. Identification of motion with echo state network [ J ]. OCEANS '04. MTS/IEEE TECHNO-OCEAN '04. 2004, 1205 ~ 1210.

[ 14 ] Skowronski MD, Harri JC. Automatic speech recognition using a predictive echo state network classifier [ J ]. Neural Networks, 2007, 414 ~ 423.

[ 15 ] Lin X, Yang Z, Song Y. Short-term stock price prediction based on echo state networks [ J ]. Expert Systems with Applications, 2009, 7313 ~ 7317.

[ 16 ] 韩敏,穆大芸. 基于贝叶斯回归的多核回声状态网络研究 [ J ]. 控制与决策,2010,25(4):531 ~ 541.

[ 17 ] Muller SD, Marchetto J, Airaghi S, et al. Optimization based on bacterial chemotaxis [ J ]. IEEE Transaction of Evolutionary Computation, 2002, 6(1): 16 ~ 29.

[ 18 ] 李威武,王慧,邹志君,钱积新. 基于细菌群体趋药性的函数优化方法 [ J ]. 电路与系统学报,2005,10(1):58 ~ 63.

[ 19 ] Zhang Y, Wu L. Stock market prediction of S&P 500 via combination of improved BCO approach and BP neural network [ J ]. Expert Systems with Applications, 2009, 8849 ~ 8854.

[ 20 ] Kaboudan MA. Genetic programming prediction of stock prices [ J ]. Computational Economics, 2000(16): 207 ~ 236.

[ 21 ] 史志伟,韩敏. ESN 岭回归学习算法及混沌时间序列预测 [ J ]. 控制与决策,2007,22(3):258 ~ 261.

## Stock Index Prediction Based on Optimized Echo State Network

ZHANG Ning-zhi, ZHOU Jia-li, SUN Wu-jun

(School of Business, Nanjing University, Nanjing 210093, China)

**Abstract:** Stock market analysis and forecasting have been the focus of economic research. Many research methods, such as fundamental analysis, technical analysis and time series analysis, have been introduced in the existing literature. Computational Intelligence is a set of nature-inspired computational methodologies, which incorporate the laws of nature into the optimization process through computer programs. Neural network has become one of primary computational intelligence means for price prediction in recent years. As a novel approach to recurrent neural network training, echo state network (ESN) has better chaotic series predictability and higher convergence rate than traditional networks. However, choosing initial transient is still based on experience rather than reliable method.

The second part introduces ESN structure and bacterial colony chemotaxis (BCC) algorithm, and compares ESN with other neural network algorithms. We further introduce BCC-ESN structure and algorithm, discuss our proposed model using the BCC algorithm to optimize the weight structure of the ESN. The third part is a simulation experiment, using the BCC-ESN model to predict stock price. The purpose of this experiment is to analyze the prediction accuracy and efficiency under different circumstances. The experimental results show that BCC algorithm on the ESN optimization can improve prediction accuracy.

The purposes of this paper are to address two primary issues. Firstly, due to the special structure of ESN reserve pool the training process is much more time-consuming when using BCC algorithm to optimize ESN. This is because a large number of weights need to be optimized by using the BCC algorithm. Secondly, the generalization ability is a very important aspect of neural network learning. When the BCC-ESN structure is too large, the training process may be faced with over-fitting problems, thereby making the fitness more than the optimal fit theory. Over-fitting model actually contains wrong information, and the predictive ability is very poor. Reflected in the specific experiment, the middle layer node is selected in the training process. The paper constructs the BCC-ESN model, which is likely to face the over-fitting problem. However, we can achieve effective solution as long as the number of nodes of the middle layer can be limited in a reasonable range.

In conclusion, this paper reviews basic ESN ideas and their algorithms. A new BCC-ESN model optimizing ESN structure with BCC algorithm is used to predict the closing price of Shanghai Security Index. The experiments show that BCC-ESN model has more predictive accuracy compared with BP network, Elman network and ESN.

**Key words:** echo state network; bacterial colony chemotaxis; stock index prediction; algorithm optimization

中文编辑: 杜 健; 英文编辑: Charlie C. Chen

---

(上接第 73 页)

matrix into Ucinet. Thirdly, the matrix was transformed into a 0-1 matrix in order to calculate the degree, freeman betweenness, and closeness for each node. Finally we tried to find whether there are some cooperative groups within the management discipline co-authorship network using K-core analysis.

Our research results show that some authors, such as Eric T. Bradlow, and Steven M. Shuganowned have close relationships with other authors, and play central roles and core positions in the co-author network. The results demonstrate that there were fewer opportunities of having cooperation among authors when the co-author network is loosely organized. In addition, some author groups have close connections among group members when they are controlled by coercive force. Based on the K-core analysis, 11 cooperative groups are found and the largest cooperative group is centered on topics, such as supply chain management, operations research, product management. The interdisciplinary study is primarily related to management science, operations management and marketing.

**Key words:** management discipline; social network analysis; co-authorship analysis

中文编辑: 杜 健; 英文编辑: Charlie C. Chen