

文章编号: 1005-8451 (2011) 11-0004-04

## 基于改进 RBF 神经网络的产业损害预警指数预报

郭恒川, 任波

(洛阳理工学院 计算机与信息工程系, 洛阳 471023)

**摘要:** 为了有效的保护产业安全, 产业损害预警指数的预报成为重要的研究方向。针对这种非线性时间序列和产业损害预警系统的应用特点, 本文对 RBF 网络的学习算法进行了一定的改进, 提高了预测结果的稳定性。根据实验仿真结果显示, 该模型优于传统使用的分析方法, 为各生产行业的生产计划提供决策支持。

**关键词:** RBF; 神经网络; 产业损害预警指数; 仿真

**中图分类号:** TP391 **文献标识码:** A

### Forecast of Industry Injury Early-warning Index based on improved RBF Neural Networks

GUO Heng-chuan, REN Bo

(Department of Computer and Information Engineering, Luoyang Institute of Science and Technology, Luoyang 471023, China)

**Abstract:** In order to protect industry security effectively, the forecast of Industry Injury Early-warning Index became an important research direction. In view of this nonlinear time series and the characteristics of Industry Injury Early-warning System, it was improved the algorithm of RBF Neural Networks, increased the stability of forecast. The simulation results of the experiment showed that the model was superior to the traditional methods, and could provide decision support for the production plant with various enterprises.

**Key words:** RBF; Neural Network; Industry Injury Early-warning Index; simulation

为实现产业保护的前置化, 我国建立并不断完善了产业损害预警监测机制<sup>[1]</sup>。各个行业和区域都开始研究并开发出了产业损害预警系统, 并取得了一些良好的使用效果。但多数系统中使用的传统预测方法仍有很多不足。而人工神经网络由于其自组织、自学习能力、非线性逼近能力, 将它用于时间序列预测更为有效, 它可以对网络进行反复的训练, 从而使网络学习存贮的知识用于推理时更接近实际系统的值<sup>[2]</sup>。

在神经网络算法中, RBF (径向基函数) 神经网络是一种局部逼近网络, 它在函数逼近能力、分类能力和学习速度等方面均优于当前应用最广泛的 BP 神经网络。将它应用于产业损害预警系统的预测会有很好的效果, 当使用 RBF 神经网络预测出的产业损害预警指数处于危险等级时, 可及时提醒厂家调整生产计划, 规避风险, 保证产业安全。

### 1 RBF 网络结构

RBF 神经网络是一种 3 层前馈网络, 网络中的神经元分层排列, 分别是输入层、隐含层和输出层, 其中每个神经元只与前一层神经元相连。输入层节点只传递输入信号到隐含层, 隐含层节点使用局部指数衰减的非线性函数 (如高斯函数) 对非线性输入—输出映射进行局部逼近, 而输出节点是简单的线性函数。当隐含层节点数足够多时, RBF 神经网络能够逼近任意的非线性函数, 因此它能成功的用于非线性函数逼近、时间序列分析、模式识别等<sup>[3]</sup>。RBF 神经网络的目标是通过计算出的历史产业损害预警指数预测下一时间 (季度或年) 的指数。RBF 神经网络结构如图 1。其结构是一个具有 N 个输入节点, P 个隐节点, M 个输出节点的网络。其中  $x$  是网络输入向量, 第  $i$  个隐含层神经元的输出为  $\varphi(\|x-c\|)$ , 式中的  $\varphi$  为径向基函数。y 为网络的输出, 输出层神经元采用线性激活函数。隐含层节点中的基函数对输入信号将在局部产生响应, 常用的基函数是高斯函数:

收稿日期: 2010-12-30

基金项目: 河南省产业损害预警系统研究 (0513053700);

洛阳理工学院青年基金项目 (2010QZ21)。

作者简介: 郭恒川, 讲师, 任波, 助教。

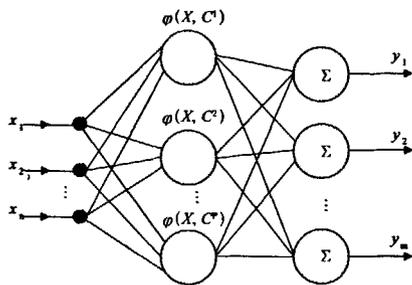


图1 RBF神经网络结构图

$$\varphi(x) = \exp\left(-\frac{\|x - c^i\|^2}{2\sigma^2}\right), i = 1, 2, \dots, P \quad (1)$$

式中  $c^i$  是第  $i$  个基函数的中心,  $\sigma_i$  是第  $i$  个感知的变量, 它决定了基函数中心点的宽度;  $\|x - c^i\|$  是  $x$  和  $c^i$  之间的欧式距离。采用高斯基函数具有的优点: (1) 表示形式简单; (2) 径向对称; (3) 光滑性好, 任意阶导数均存在; (4) 便于进行理论分析<sup>[4]</sup>。输出层对隐节点的输出进行线性加权组合, 即:

$$y_k = \sum_{i=1}^P w_{ik} \varphi_i(x), k = 1, 2, \dots, M \quad (2)$$

式中  $M$  是输出节点数,  $w_{ik}$  是输出节点  $k$  与隐层节点  $i$  之间的权值。

## 2 RBF网络的学习算法

RBF网络学习算法由2个阶段组成: (1) 根据输入的样本为隐含层各节点的径向基函数确定合适的数据中心  $c^i$ , 并根据各中心之间的距离确定隐节点的宽度  $\sigma_i$ 。(2) 监督学习阶段, 用于监督学习算法训练输出层的权值。一旦确定了径向基函数的中心  $c^i$ , 所有的训练样本  $\varphi_i$  和预期输出  $y_k$  是已知的, 输出权值  $w_j$  可以由LMS算法或伪逆法等方法求出。因此, RBF网络的学习算法主要解决的问题是如何确定各径向基函数的数据中心。

### 2.1 数据中心确定的常用方法

确定RBF网络数据中心的方法主要有以下几种: (1) 固定法; (2) 随机选取法; (3) K-均值聚类法; (4) 正交最小二乘(OLS)法。

这些常用算法虽然收敛速度比较快, 但是其结果与初始聚类中心有关, 稳定性难以保证。

### 2.2 K-均值算法与蚁群算法的混合算法

蚁群聚类算法是基于生物蚁群系统的集体觅

食行为而发展起来的一类仿生化算法。蚁群算法初始化时, 很难在短时间内从大量的路径中找到一条较好的路径, 所以收敛速度较慢。因此, 改进的RBF神经网络在中心数据选择时先用K-均值算法作快速分类, 然后再根据分类结果更新信息素, 指导其他蚂蚁选择, 可以加快收敛速度。改进后的混合算法具体操作步骤如下:

(1) 初始化基本参数: 样本个数  $N$ , 样本属性  $m$ , 聚类半径  $r$ , 允许误差值  $\varepsilon_0$ 。

(2) 使用K-均值算法对输入样本进行快速分类。

(3) 采用欧式距离计算任意样本直接的距离  $d_{ij}$ 。

(4) 如果  $d_{ij} \leq r$  对其信息素  $\tau_{ij}$  置1, 否则  $\tau_{ij}$  置0。

(5) 根据K-均值算法分类结果, 计算出聚类中心  $z_j$ , 计算样本  $x_i$  到聚类中心  $z_j$  的距离  $d_{ij}$ 。若样本  $x_i$  分配到聚类中心  $z_j$ , 则修改信息素为  $\tau_{ij} + Q / d_{ij}$ , 其他路径不变。否则, 所有路径的信息素增量  $\tau_{ij} + Q / d_{ij}$ 。

(6) 根据公式计算  $x_i$  聚类到  $x_j$  的概率  $P_{ij}$ 。若  $P_{ij}(t) \geq P_0$ , 则  $x_i$  归属于  $x_j$  类; 否则分为2类。

(7) 计算能和  $x_j$  归一类的聚类中心  $c_j$ , 计算总体误差  $\varepsilon$ , 若  $\varepsilon > \varepsilon_0$ , 则继续执行, 如果满足误差要求则停止算法。

(8) 计算各样本到新的聚类中心之间的距离  $d_{ij}$ , 并修改信息素为  $\tau_{ij} + Q / (1 + d_{ij})$ 。

(9) 重复(6)~(8), 直到满足误差, 终止算法, 得到聚类个数  $K$ , 以及  $K$  个聚类中心。

完成上述混合算法后, 获得RBF神经网的隐层节点个数以及径向基函数的数据中心。

## 3 网络构造与测试

### 3.1 样本选取与处理

本文提出的神经网络预测模型是对某省的产业损害预警指数进行的预测, 该指数可以看作一个有  $L$  个元素的时间序列来进行处理。通过序列前  $N$  个时刻的值预测出后  $M$  个时刻的值, 可以将数据分为  $k$  个长度为  $N+M$ 、有一定重叠的数据段样本, 得到  $k=L-(N+M)+1$  个样本<sup>[5]</sup>。每个样本的前  $N$  个值作为RBF神经网络的输入, 后  $M$  个值作为目标输出。通过网络学习, 实现从输入到输出的映射, 达到时间序列预测的目的。由于产业损害预测指数取值在0和1之间, 因此无需进行归一化处

理。实践中所取的样本数据为该省2007年1季度到2009年4季度的氧化铝产业损害指数，共计12个样本。将每连续3个季度的指数作为RBF网络模型的输入，输出为预测的第4个季度的指数，即输入层神经元为3个，输出层神经元为1个。这样，数据共分为9组，如表1所示。其中，前7组用于训练，后2组用于检验。

表1 样本数据与划分方法

3个输入	1个输出
2007年1季度~2007年3季度 0.327、0.415、0.463	2007年4季度 0.384
2007年2季度~2007年4季度 0.415、0.463、0.384	2008年1季度 0.421
2007年3季度~2008年1季度 0.463、0.384、0.421	2008年2季度 0.434
2007年4季度~2008年2季度 0.384、0.421、0.434	2008年3季度 0.482
2008年1季度~2008年3季度 0.421、0.434、0.482	2008年4季度 0.460
2008年2季度~2008年4季度 0.434、0.482、0.460	2009年1季度 0.465
2008年3季度~2009年1季度 0.482、0.460、0.465	2009年2季度 0.422
2008年4季度~2009年2季度 0.460、0.465、0.422	2009年3季度 0.380
2009年1季度~2009年3季度 0.465、0.422、0.380	2009年4季度 0.254

### 3.2 基于matlab工具箱的网络设计

根据上述分析的网络模型对产业损害预警指数进行预测，其创建采用matlab径向基函数神经网络工具箱 Net=newrb(P, T, GOAL, SPREAD) 来实现。函数中P为输入向量；T为目标向量；SPREAD为径向基函数的分布密度，SPREAD的取值越大，则网络的预测性能越平滑，考虑到分布密度过大可能导致计算上出现问题，在这里设置成默认值1。隐层单元激励函数采用radbas，加权函数采用dist，输入函数采用netprod，输出层神经元激励函数采用纯线性函数purelin，加权函数为dotprod，输入函数为netsum。

### 3.3 预测结果与分析

训练前需要对网络参数进行初始化，不同的参数影响其收敛速度与预测结果的精度，经过大量的实验即可得到最优的参数应用于网络。图2中用改进的RBF神经网络进行预测，比产业损害预警系统中所用的传统三次曲线趋势外推法，拟合精度高，学习速度快，总体预测效果令人满意。

2009年度各季度的输出值与实际值之间的拟合如表2所示。该网络训练的输出值和实际值之间的差值较小，且非常稳定，平均误差率仅为6.41%。说明该网络能较好的应用于产业损害预警指数的预测，对生产计划的指导有实际的应用价值。

由于产业损害预警指数受到国内外政治，经济环境的影响较大，因此各种预测方法的误差率

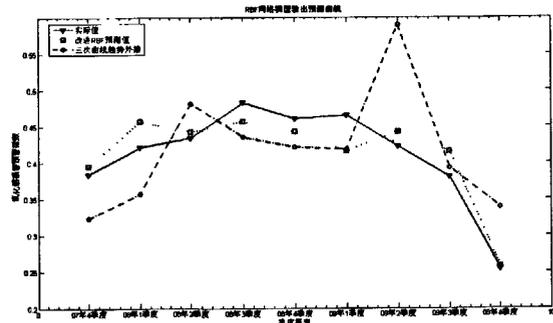


图2 预测曲线和真实数据的比较图

表2 预测数据与实际数据比较

时间	09年4季度	09年3季度	09年2季度	09年1季度	平均误差率
实际值	0.254	0.380	0.422	0.465	
改进型RBF神经网络	0.257	0.415	0.442	0.416	6.41%
三次曲线趋势外推法	0.339	0.393	0.589	0.418	21.63%
RBF误差率	1.18%	9.21%	4.73%	10.53%	6.41%
外推法误差率	33.46%	3.4%	39.57%	10.10%	21.63%

均大于应用在受人为因素影响较小的方面，如降雨量的预测。如果采用更大容量的训练样本，如收集并计算出最近20年各个产业损害预警指数，就会获得更好的预测性能。

另外，目前该模型每次只有预测下一季度的产业损害指数，为预测下N（大于1）个月的损害指数，可以把历史数据加上被预测的N-1个季度的数据作为输入样本，对未来第二季度、第三季度进行预测。这样不仅可以提高效率，还能增加实用性。这样的滚动预测，可以使生产计划的制定更具有前瞻性。

## 4 结束语

本文提出了一种将改进的蚁群算法和RBF神经网络学习算法相结合的神经网络预测模型，用蚁群算法良好的全局搜索能力来弥补神经网络在全局搜索上的缺陷，同时对蚁群算法的收敛速度慢的缺点进行了相应改进。实践证明这种神经网络预测模型比传统产业损害预测模型训练结果精度高，预测结果平稳，取得了较好的实际效果。这种模型同时可推广到各种时间序列数据预测的领域，有一定的普遍实用性。

### 参考文献：

[1] 郭恒川，任波. Web环境下产业损害预警图的绘制与显

- 示[J]. 甘肃联合大学学报, 2007, 21 (6): 66-69.
- [2] 张玉瑞, 陈剑波. 基于RBF神经网络的时间序列预测[J]. 计算机工程与应用, 2005 (11): 74-76.
- [3] 施彦, 韩力群, 廉小亲. 神经网络设计方法与实例分析[M]. 北京: 邮电大学出版社, 2009.

- [4] 王静, 田丽, 蒋慧. 基于遗传算法的RBF网络的短期电力负荷预测[J]. 电子技术, 2010 (4): 15-16.
- [5] 何迎生, 段明秀. 基于RBF神经网络的时间序列预测[J]. 吉首大学学报(自然科学版), 2008, 29 (3): 52-54.

责任编辑 陈蓉

(上接 P3)

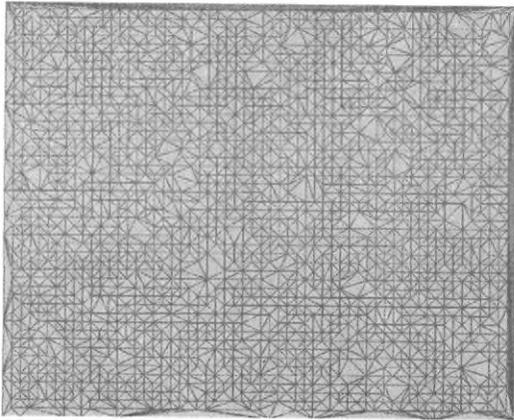


图4 SRTM生成模型俯视图

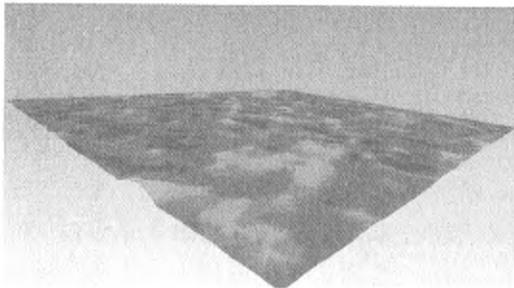


图5 ASTER-GDEM生成模型侧视图

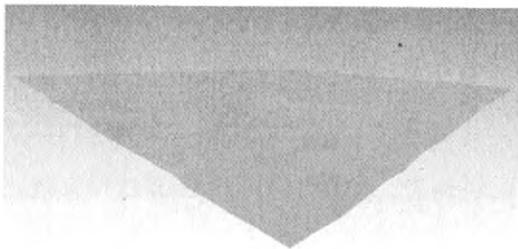


图6 SRTM生成模型侧视图

将地形模型进行贴图的优势是可以方便的将线路进行定位, 并且在布置周边景物时, 便于确定周边场景的地物位置和地物类型。使用 ASTER-GDEM 高程数据, 将采样间隔定为 30 m, 并进行贴图后得到的地形模型如图 7。

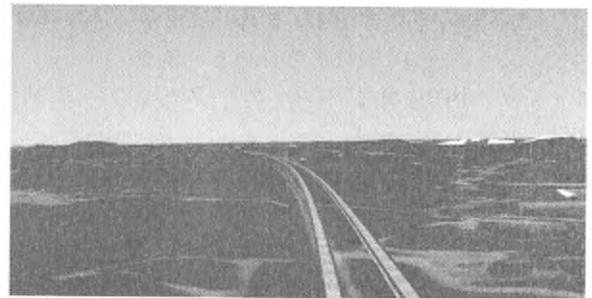


图7 地形模型最终效果图

## 9 结束语

本文以上海磁浮线路周边地形作为仿真对象, 对利用 Multigen Creator terrain pro 组件生成地形仿真模型的各个步骤进行了较为详细的分析, 并通过对各环节进行对比, 提出优化方案, 进而提升了模型的精确度和仿真度。列车线路周边实景地形仿真模型不仅可以应用在轨道交通驾驶仿真领域, 而且对于军事领域的相关应用也有着较高的参考价值, 因此本文将有助于进一步完善利用 Multigen Creator 软件平台生成实景地形模型的一整套方法。

参考文献:

- [1] Multigen-Paradigm, Inc. Creating Terrain for Simulations[M]. USA: Multigen-Paradigm, Inc, 2004.
- [2] 王乘, 周均清, 李利军. Creator 可视化仿真建模技术[M]. 武汉: 华中科技大学出版社, 2005.
- [3] 张祖勋, 张剑清. 城市建模的途径与关键技术[J]. 世界科技研究与发展, 2003, 25 (6): 23-28.
- [4] 丁斌, 祖家奎. 三维可视化中的地形建模与实现技术研究[J]. 计算机仿真, 2008, 25 (12): 212-215.
- [5] 张寅宝. 三维地形多分辨率建模与精度评估[D]. 郑州: 解放军信息工程大学硕士论文, 2006.
- [6] 朱婷婷, 翟婉明, 赵春发. 磁悬浮线路景观三维几何建模[J]. 交通运输工程与信息学报, 2006 (1): 56-59.

责任编辑 陈蓉