

DOI: 10.11992/tis.201608009

网络出版地址: <http://kns.cnki.net/kcms/detail/23.1538.TP.20170407.1758.014.html>

基于相对贡献指标的自组织 RBF 神经网络的设计

乔俊飞^{1,2}, 安茹^{1,2}, 韩红桂^{1,2}

(1. 北京工业大学 电子信息与控制工程学院, 北京 100124; 2. 计算智能与智能系统北京市重点实验室, 北京 100124)

摘 要: 针对 RBF(radial basis function)神经网络的结构和参数设计问题, 本文提出了一种基于相对贡献指标的自组织 RBF 神经网络的设计方法。首先, 提出一种基于相对贡献指标 (relative contribution, RC) 的网络结构设计方法, 利用隐含层输出对网络输出的相对贡献来判断是否增加或删除 RBF 网络相应的隐含层节点, 并且对神经网络结构调整过程的收敛性进行证明。其次, 采用改进的 LM (Levenberg-Marquardt algorithm) 算法对调整后的网络参数进行更新, 使网络具有较少的训练时间和较快的收敛速度。最后, 对提出的设计方法进行非线性函数仿真和污水处理出水参数氨氮建模, 仿真结果表明, RBF 神经网络能够根据研究对象自适应地动态调整 RBF 结构和参数, 具有较好的逼近能力和更高的预测精度。

关键词: RBF 神经网络; 相对贡献指标; 改进的 LM 算法; 结构设计; 出水氨氮; 收敛速度; 预测精度

中图分类号: TP183 **文献标志码:** A **文章编号:** 1673-4785(2018)02-0159-09

中文引用格式: 乔俊飞, 安茹, 韩红桂. 基于相对贡献指标的自组织 RBF 神经网络的设计 [J]. 智能系统学报, 2018, 13(2): 159-167.

英文引用格式: QIAO Junfei, AN Ru, HAN Honggui. Design of self-organizing RBF neural network based on relative contribution index[J]. CAAI transactions on intelligent systems, 2018, 13(2): 159-167.

Design of self-organizing RBF neural network based on relative contribution index

QIAO Junfei^{1,2}, AN Ru^{1,2}, HAN Honggui^{1,2}

(1. College of Electronic Information and Control Engineering, Beijing University of Technology, Beijing 100124, China; 2. Beijing Key Laboratory of Computation Intelligence and Intelligence System, Beijing 100124, China)

Abstract: A design method for a self-organizing RBF Neural Network based on the Relative Contribution index is proposed with the aim of performing the structural design and parameter optimization of the Radial Basis Function (RBF) neural network. First, a self-organizing RBF network design method based on the Relative Contribution (RC) index is proposed. The relative contribution of the output of the hidden layer to the network output was used in order to assess whether a node of the hidden layer corresponding to the RBF network was inserted or pruned. Additionally, the convergence of the adjustment process of the neural structure was proven. Secondly, the adjusted network parameters were updated by the improved Levenberg-Marquardt (LM) algorithm in order to reduce the training time and increase the convergence speed of the network. Finally, the proposed algorithm was used in the simulation of the nonlinear function, and the modeling of the ammonia and nitrogen sewage effluent parameters. The simulation results revealed that the structure and parameters of the RBF neural network could be adjusted adaptively and dynamically according to the object under investigation, and that they had excellent approximation ability and higher prediction accuracy.

Keywords: RBF neural network; relative contribution index; improved LM algorithm; structure design; ammonia and nitrogen effluent parameters; convergence speed; prediction accuracy

收稿日期: 2016-08-29. 网络出版日期: 2017-04-07.

基金项目: 国家自然科学基金重点项目 (61533002, 61225016);
北京市教育委员会科研计划项目 (km201410005002);
高等学校博士学科点基金项目 (20131103110016).

通信作者: 乔俊飞. E-mail: anru@emails.bjut.edu.cn.

径向基函数 (RBF) 神经网络由于其准确的局部感知特性和强大的非线性逼近能力, 已被广泛应用于时间序列预测、非线性系统在线建模与控

制等领域^[1]。神经网络的研究和应用对于复杂工业过程建模提供了新的方法;与此同时,神经网络的设计成为非线性动态过程建模的重点和难点。RBF神经网络的性能严重依赖于网络的结构和参数,包括隐含层节点的数量、中心和宽度、隐含层到输出层的连接权值等参数^[2]。事实上,如果网络结构越大,系统动力学错综复杂,表现良好的性能,但是会增大计算量和计算时间,容易导致“过拟合”;反之,规模过小,内部动力学过于简单,学习能力不足,不能高精度地逼近非线性映射,而且网络的参数优化算法能够保证网络达到较好的性能,能够较快收敛,因此,RBF网络的结构和参数学习算法研究对于实际应用和理论研究具有重要的实际意义。

为使RBF神经网络能够根据实际被控过程自适应的动态调整,近年来不少学者进行研究。Platt^[3]首先提出资源分配网络算法(resource allocation network, RAN),根据实际对象能够动态增加RBF网络隐含层节点,但是该算法只增加隐节点不能删减,对于处理复杂问题时网络结构会出现冗余。在此基础上,为解决这一问题,Lu等^[4]提出一种最小资源分配算法(minimal resource allocation network, MRAN),该算法能够在学习过程中增加和删减神经元动态调整去适应被控过程,该方法受到广泛应用,但是却忽略了神经元结构调整之后的参数学习,导致网络收敛速度较慢^[5]。Huang^[6]提出一种在线调整的RBF结构设计方法;随后出现广义增长修剪算法(generalized growing and pruning RBF, GGAP-RBF),根据隐含层神经元的重要性判断是否增删,但仅仅考虑对新增加或者删减的神经元进行参数调整,有效提高算法的运算速度,但是网络初始值的设定需要依据整体样本数据,因而不适合在线学习。Gonzales^[7]采用进化计算的思想,利用其良好的鲁棒性和全局搜索能力对网络的结构和参数进行调整学习,取得较好的效果。文献[8-9]提出用粒子群(particle swarm optimization, PSO)算法自动调整每个RBF网络的中心,宽度和权值,能够获得不错的建模性能,但是该算法由于训练过程中需要全局搜索,需要较长的训练时间而且算法计算复杂,不利于实时在线建模。Lian^[10]提出自组织RBF网络(self-organizing RBF, SORBF),设计仅仅采用训练误差作为结构调整判断条件,没有考虑隐含层和输出层之间的相关性以及网络调整后参数设置问题,训练时间较长。Yu等^[11]提出基于误差修正的思想,每次采用误差最大的点对应的输入数据作为新增加神经元的中心,采用改进的LM算法优化所有隐节点的参数,训练时间快,能够采用精简的结构去逼近非线性函数,但是该自

组织机制只能增加隐节点不能删减,而且对于多输入的输入数据如何判断中心不能很好地确定。要想使RBF网络具有更好的非线性建模性能,完成结构设计之后,寻找快速的参数学习算法对网络的性能也是至关重要的。常见的参数训练算法有BP算法、高斯-牛顿算法和LM算法等。其中,最常用的是BP算法,但是该算法由于搜索空间限制容易陷入局部极小,收敛速度慢等。近年来,具有快速收敛速度和强有力搜索空间的二阶LM算法被引入到训练RBF网络,取得了很好的效果,获得了广泛应用^[12-13]。但是该算法由于雅可比矩阵的计算会增加计算量和存储空间,影响算法的训练速度,因此,本文采用一种改进的LM算法对参数进行训练。

针对RBF网络的结构和参数设计问题,首先,利用隐含层和输出层之间的回归关系,采用相对贡献指标结合训练误差信息处理能力,最大程度地挖掘隐含层和输出层之间的隐含信息,解决了RBF网络结构动态调整的问题;同时,采用改进的LM算法,将整个雅可比矩阵的计算转化为向量相乘的形式,避免整个雅可比矩阵的存储,加速算法的训练时间和收敛速度;最终保证RBF网络能够根据实际处理对象的动态过程快速准确地自适应调整结构和参数,达到满意的非线性逼近能力和预测精度,验证了算法的有效性。

1 RBF 网络

RBF网络是一个包括输入层、隐含层和输出层的3层前馈神经网络,其拓扑结构如图1所示。

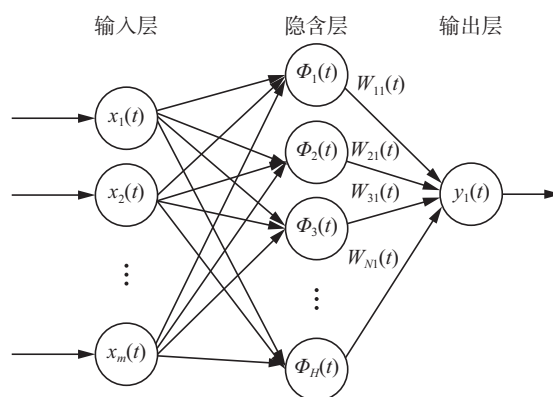


图1 RBF网络拓扑结构图

Fig. 1 The topology structure of RBF network

图1中网络的输入向量为

$$\mathbf{x}(t) = [x_1(t) \ x_2(t) \ \cdots \ x_M(t)]^T,$$

M 为输入向量的维数,隐含层激活函数采用高斯函数,隐含层第 j 个神经元的输出表达式如式(1)所示。

$$\Phi_j(t) = \exp\left(-\frac{\|\mathbf{x}(t) - \mathbf{c}_j(t)\|^2}{\sigma_j^2(t)}\right) \quad (1)$$

式中: $\Phi_j(t)$ 为隐含层第 j 个神经元的输出; $\|\cdot\|$ 为欧几里得范数; $\mathbf{c}_j(t) = [c_{j1}(t) \ c_{j2}(t) \ \cdots \ c_{jH}(t)]^T$, $j = 1, 2, \dots$; H 为径向基函数的中心向量, H 为隐含层神经元的个数, 隐节点中心离输入越近, 输出值越大; $\sigma_j(t)$ 为第 j 个基函数的扩展宽度, 主要影响基函数的分布的密集程度^[14]。输出层神经元的输出为

$$y_1(t) = \sum_{j=1}^H \mathbf{w}_{j1} \Phi_j(t) \quad (2)$$

式中: \mathbf{w}_{ji} 为第 j 个隐含层神经元与输出层的连接权值向量, $y_1(t)$ 为输出层神经元的输出。

2 自组织 RBF 网络的设计

2.1 相对贡献指标

网络结构调整的设计思想是采用回归的思想, 分别对隐含层和输出层矩阵进行成分提取为 t_i 和 v_i , 要求提取的成分对于原变量信息具有最大的解释能力, 而且具有最大的相关性, 得到隐含层神经元和输出层神经元的相对贡献指标, 用来表示此神经元对输出神经元的贡献程度; 同时结合网络的误差信息处理能力, 作为判断结构是否调整的依据。

相对贡献指标: RBF 神经网络隐节点 j 在样本数为 P 个的相对贡献指标定义为

$$RC(j) = \frac{\sum_{i=1}^r (\mathbf{Y}_{ij} \beta_i)}{\sum_{h=1}^H (\sum_{i=1}^r \mathbf{Y}_{ih} \beta_i)} \quad (3)$$

式中: $\mathbf{Y}_i = (\mathbf{Y}_{i1}, \mathbf{Y}_{i2}, \dots, \mathbf{Y}_{iH})$ 代表提取的成分 t_i 的权重向量; β_i 为成分的负荷量; H 为隐含层节点的数量; r 为成分的数量, 每次提取的成分按照式 (4)~(8) 计算。

$$\mathbf{t}_i = \Phi_{i-1} \mathbf{Y}_i, 0 \leq i \leq r \quad (4)$$

$$\Phi_{i-1} = \mathbf{t}_i \alpha_i^T + \Phi_i \quad (5)$$

$$\mathbf{v}_i = \mathbf{y}_{i-1} \mathbf{u}_i \quad (6)$$

$$\mathbf{y}_{i-1} = \mathbf{v}_i \beta_i^T + \mathbf{y}_i \quad (7)$$

$$\beta_i = \frac{\mathbf{y}_i^T \mathbf{t}_i}{\|\mathbf{t}_i\|^2} \quad \alpha_i = \frac{\Phi_i^T \mathbf{t}_i}{\|\mathbf{t}_i\|^2} \quad (8)$$

式中: \mathbf{t}_i 和 \mathbf{v}_i 是第 i 对成分的得分向量; α_i 和 β_i 为成分的负荷量; Φ_i 和 \mathbf{y}_i 为第 i 次迭代产生的残差矩阵, 不断迭代直到得到 r 个成分, 并且 $r = \text{rank}(\Phi_0)$; Φ_0 是隐含层矩阵和输出矩阵对 P 个样本的标准化矩阵。另外要求提取的成分具有最大的相关性, 通过拉格朗日法转化为求权重向量 \mathbf{Y}_i 和 \mathbf{u}_i 。进而求出相对贡献指标。

2.2 结构调整机制

2.2.1 神经元增加机制

结合上述相对贡献指标 RC 和网络误差信息处理能力作为神经元增加机制的判定条件, 当算

法迭代 t 次的误差比 $t-n$ 大时, 代表此时网络对于动态过程的信息处理能力不足, 需要增加新的隐节点, 分裂当前隐含层神经元和输出神经元之间具有最大相对贡献的隐含层神经元 j , 此神经元表示和输出神经元之间具有最大的贡献度, 即满足式 (9):

$$\begin{cases} E(t) - E(t-n) > 0 \\ j = \arg \max_{1 \leq j \leq H} (RC(j)) \end{cases} \quad (9)$$

式中: $E(t)$ 和 $E(t-n)$ 分别为迭代步数 t 和 $t-n$ 时的训练误差, n 是样本间隔, j 是隐含层神经元和输出神经元相对贡献 RC 最大的神经元, H 是在 t 时刻存在的隐含层神经元数量。对新增加的神经元参数设置为

$$\mathbf{c}_{\text{new}}(t) = \frac{1}{2} (\mathbf{c}_j(t) + \mathbf{x}(t)) \quad (10)$$

$$\sigma_{\text{new}}(t) = \sigma_j(t) \quad (11)$$

$$\mathbf{w}_{\text{new}}(t) = \mathbf{e}_j(t) (\Phi_{\text{new}}(t))^{-1} \quad (12)$$

式中: $\mathbf{c}_j(t)$ 、 $\sigma_j(t)$ 分别代表第 j 个神经元分裂前的中心和宽度; \mathbf{c}_{new} 、 σ_{new} 代表新增加神经元的中心和宽度; \mathbf{w}_{new} 为新增加神经元的输出连接权值; $\mathbf{e}_j(t)$ 为 t 时刻神经网络的误差; $\Phi_{\text{new}}(t)$ 为新增加神经元的隐含层输出值。

2.2.2 神经元删减机制

如果当前第 k 个隐含层神经元和输出层的 RC 小于设定的阈值 ε , 说明此神经元对输出的相对贡献较小, 可以认为此神经元对输出的贡献很小甚至可以忽略, 所以删掉第 k 个神经元, 则满足:

$$\begin{cases} \min(RC(j)) < \varepsilon \\ k = \arg \min_{1 \leq j \leq H} (RC(j)) \end{cases} \quad (13)$$

式中: k 是隐含层神经元和输出神经元相对贡献最小的神经元, ε 为设定的删减阈值。

$$\mathbf{c}'_{i-1}(t) = \mathbf{c}_k(t) \quad (14)$$

$$\sigma'_{i-1}(t) = \sigma_k(t) \quad (15)$$

$$\mathbf{w}''_{i-1}(t) = \mathbf{w}'_{i-1}(t) + \frac{\Phi_k(t)}{\Phi'_{i-1}(t)} \mathbf{w}_k(t) \quad (16)$$

式中: $\mathbf{c}'_{i-1}(t)$ 和 $\sigma'_{i-1}(t)$ 分别为与隐含层神经元 i 欧式距离最近的隐含层神经元的中心、宽度; $\mathbf{c}_k(t)$ 和 $\sigma_k(t)$ 分别为第 k 个隐含层神经元的中心和宽度; $\mathbf{w}'_{i-1}(t)$ 和 $\mathbf{w}''_{i-1}(t)$ 分别为删除第 k 个神经元之前与隐含层神经元 k 欧式距离最近的隐含层神经元的连接权值和删除之后此神经元的连接权值; $\Phi_k(t)$ 和 $\Phi'_{i-1}(t)$ 分别第 k 个隐含层神经元的输出值和为删除此神经元之前, 与隐含层神经元 i 欧式距离最近的隐含层神经元的输出值。

2.3 改进的 LM 算法

采用改进的 LM 算法 (improved LM algorithm, ILM) 去优化 RBF 网络的中心、宽度和权值。不

同于之前算法的矩阵运算方式,文中采用一种将矩阵运算转化为向量相乘的方式,不需要存储整个雅可比矩阵,这样避免整个雅可比矩阵的计算,适用于输入数量比较多。其中,拟海森矩阵和梯度向量的计算通过子矩阵及其子向量叠加和的方式来得到的^[15]。同时,在训练过程中算法引入自适应学习率,这样能够加快网络的收敛速度,提高算法的预测精度。具体更新规则如式(17)所示:

$$\mathbf{\Omega}(t) = [w_1(t) \cdots w_h(t) \cdots w_H(t) \ c_{11}(t) \cdots c_{1,j}(t) \cdots c_{h,p}(t) \ \sigma_1(t) \cdots \sigma_h(t) \cdots \sigma_H(t)] \quad (19)$$

拟海森矩阵 $\mathbf{Q}(t)$ 和梯度向量 $\mathbf{\Theta}(t)$ 的计算分别是对应的子矩阵与子向量的累加和得到的,计算公式为

$$\mathbf{Q}(t) = \sum_{p=1}^P \mathbf{q}_p(t), p=1, 2, \cdots, P \quad (20)$$

$$\mathbf{\Theta}(t) = \sum_{p=1}^P \boldsymbol{\varphi}_p(t), p=1, 2, \cdots, P \quad (21)$$

式中: P 是样本总数,子矩阵 $\mathbf{q}_p(t)$ 与子向量

$$\mathbf{j}_p(t) = \left[\frac{\partial \mathbf{e}_p(t)}{\partial \mathbf{w}_1(t)} \cdots \frac{\partial \mathbf{e}_p(t)}{\partial \mathbf{w}_H(t)} \frac{\partial \mathbf{e}_p(t)}{\partial \sigma_1(t)} \cdots \frac{\partial \mathbf{e}_p(t)}{\partial \sigma_H(t)} \frac{\partial \mathbf{e}_p(t)}{\partial c_{1,1}(t)} \cdots \frac{\partial \mathbf{e}_p(t)}{\partial c_{h,i}(t)} \cdots \frac{\partial \mathbf{e}_p(t)}{\partial c_{H,i}(t)} \right] \quad (25)$$

根据文献[17]式(25)中误差对权值的导数,误差对宽度的导数,误差对中心的导数计算公式分别为

$$\frac{\partial \mathbf{e}_p}{\partial \mathbf{w}_h} = -\frac{\partial \mathbf{o}_p}{\partial \mathbf{w}_h} = -\boldsymbol{\varphi}_h(\mathbf{x}_p) \quad (26)$$

$$\frac{\partial \mathbf{e}_p}{\partial \sigma_h} = -\frac{\partial \mathbf{o}_{p,1}}{\partial \sigma_h} = -\frac{\partial \mathbf{o}_p}{\partial \varphi_h(\mathbf{x}_p)} \frac{\partial \varphi_h(\mathbf{x}_p)}{\partial \sigma_h} = -\frac{2\mathbf{w}_{h,1}\varphi_h(\mathbf{x}_p)\|\mathbf{x}_{p,i}-\mathbf{c}_{h,i}\|^2}{\sigma_h^3} \quad (27)$$

$$\frac{\partial \mathbf{e}_p}{\partial \mathbf{c}_{h,i}} = -\frac{\partial \mathbf{o}_p}{\partial \mathbf{c}_{h,i}} = -\frac{\partial \mathbf{o}_p}{\partial \varphi_h(\mathbf{x}_p)} \frac{\partial \varphi_h(\mathbf{x}_p)}{\partial \mathbf{c}_{h,i}} = -\frac{2\mathbf{w}_{h,m}\varphi_h(\mathbf{x}_p)(\mathbf{x}_{p,i}-\mathbf{c}_{h,i})}{\sigma_h^2} \quad (28)$$

雅可比矩阵行向量 $\mathbf{j}_p(t)$ 的计算由式(26)~(28)计算得到,对于所有输入模式,拟海森矩阵和梯度向量分别由式(20)和(21)得到,然后应用更新规则(17)对3个参数同时进行更新。采用上述ILM算法去优化RBF网络参数,加快算法的收敛速度并且提高网络的预测准确度。

RC-RBF网络的设计算法的步骤如下:

1) 初始随机给定一个RBF神经网络,输入节点与输出节点个数根据具体实验设置,隐节点个数随机设定,所有的参数随机产生在一个小范围内。

2) 对于输入样本 $\mathbf{x}(t)$,隐含层到输出层之间的权值、宽度、中心分别通过式(26)~(28)进行调整,采用固定样本个数的在线形式,然后,判断神

$$\mathbf{\Omega}(t+1) = \mathbf{\Omega}(t) - (\mathbf{Q}(t) + \lambda(t)\mathbf{I})^{-1} \mathbf{\Theta}(t) \quad (17)$$

式中: $\mathbf{\Omega}(t)$ 为更新规则包含参数向量, $\mathbf{\Theta}(t)$ 为梯度向量, $\mathbf{Q}(t)$ 为拟海森矩阵, \mathbf{I} 是为了避免拟海森矩阵奇异设置的单位矩阵,根据文献[16],自适应学习率 $\lambda(t)$ 被定义为

$$\lambda(t) = \alpha \|\mathbf{e}(t)\| + (1-\alpha) \|\mathbf{j}(t)^T \mathbf{e}(t)\| \quad (18)$$

式中: α 为正实数, $\alpha \in (0, 1)$ 。参数向量 $\mathbf{\Omega}(t)$ 包括更新网络的所有参数:连接权值 \mathbf{w} ,中心向量 \mathbf{c} ,宽度向量 σ 。

$\boldsymbol{\varphi}_p(t)$ 的计算公式分别为

$$\mathbf{q}_p(t) = \mathbf{j}_p(t)^T \mathbf{j}_p(t) \quad (22)$$

$$\boldsymbol{\varphi}_p(t) = \mathbf{j}_p(t)^T \mathbf{e}_p(t) \quad (23)$$

第 p 个样本的训练误差定义为

$$\mathbf{e}_p(t) = \mathbf{y}_p(t) - \mathbf{o}_p(t) \quad (24)$$

式中: P 为样本数, $\mathbf{e}_p(t)$ 为训练误差, $\mathbf{y}_p(t)$ 与 $\mathbf{o}_p(t)$ 是分别为第 p 个输入样本对应的网络输出与实际输出,雅克比矩阵行向量 $\mathbf{j}_p(t)$ 如式(25)所示:

经网络是否满足结构调整条件,满足条件则转向3),否则转向6)。

3) 根据式(3)计算当前存在的隐含层神经元与网络输出计算相对贡献指标RC,如果第 t 步的训练RMSE比第 $t-n$ 步大时: $\mathbf{e}(t)-\mathbf{e}(t-n)>0$ 时,则转向4);转向5),判断是否满足删减条件,否则转向6)。

4) 此时,说明神经网络信息处理能力不足,需要增加神经元,按照式(9),选出相对贡献指标最大的神经元进行分裂,根据式(10)~(12)对新增神经元的设定初始化参数。

5) 如果满足删减条件(13),则删除相对指标贡献值小于阈值 ε 的隐含层神经元,删除与此神经元的连接权值、中心和宽度,并利用式(14)~(16)对其临近的神经元的相应参数进行调整。

6) 利用改进的LM算法对神经网络的参数进行更新。

7) 满足停止条件或达到计算次数时停止计算,否则转向2)(网络结构还需调整)进行重新训练。

3 收敛性分析

网络是否收敛决定了算法的性能,因为网络在结构调整之后对其收敛性往往不能保证,以下给出结构调整阶段的收敛性证明,主要分为3部分:1) 隐节点增加阶段;2) 隐节点删减阶段;3) 隐节点数目不变的阶段。

假设:当前网络存在 J 个隐含层神经元,当前时刻的误差为 $e_j(t)$ 。

1) 当满足神经元增加条件时,分裂神经元,此时神经元数目变为 $J+1$ 个,此时网络的误差变为 $e'_{J+1}(t)$ 。新增加的神经元的参数设置按照式(12)进行设置。

$$\begin{aligned} e'_{J+1}(t) &= y_p(t) - \sum_{j=1}^{J+1} w_j(t) \Phi_j(t) = \\ y_p(t) - &\left(\sum_{j=1}^J w_j(t) \Phi_j(t) + w_{\text{new}}(t) \Phi_{\text{new}}(t) \right) = \\ e_j(t) - &\frac{e_j(t)}{\Phi_{\text{new}}(t)} \Phi_{\text{new}}(t) = 0 \end{aligned} \quad (29)$$

可以看出,隐含层新增加神经元之后,其参数设置对网络的输出误差进行了补偿,调整后误差为0,一定程度上加快算法的学习速度。

2) 当满足删减条件时,删除第 k 个神经元,此时神经元的数量变为 $J-1$,神经网络的输出误差为 $e'_{J-1}(t)$,删减之后对临近神经元的参数更新设置如式(16)。

可以看出,删掉神经元与输出之间的连接权值,中心,宽度等参数,对其邻近的神经元进行参数更新,神经元调整前后,神经网络的输出误差相等,可见结构删减对网络的误差没有产生影响。

$$\begin{aligned} e'_{J-1}(t) &= y_p(t) - \sum_{j=1}^{J-1} w_j(t) \Phi_j(t) = \\ y_p(t) - &\left(\sum_{j=1, j \neq k, j \neq i-1}^J w_j(t) \Phi_j(t) + w_{i-1}(t) \Phi_{i-1}(t) \right) = \\ y_d(t) - &\left(\sum_{j=1, j \neq k, j \neq i-1}^J w_j(t) \Phi_j(t) + \right. \\ &\left. \left(w'_{i-1}(t) + \frac{\Phi_k(t)}{\Phi'_{i-1}(t)} w_k(t) \right) \Phi'_{i-1}(t) \right) = \\ y_d(t) - &\sum_{j=1, j \neq i-1}^J w_j(t) \Phi_j(t) - w'_{i-1}(t) \Phi'_{i-1}(t) e_j(t) \end{aligned} \quad (30)$$

3) 隐节点数目不变的阶段

为了证明算法在固定神经元时算法的收敛性,定义一个 Lyapunov 函数:

$$F(\mathbf{\Omega}(t)) = \frac{1}{2} \mathbf{e}^T(t) \mathbf{e}(t) \quad (31)$$

根据泰勒展开式可以得到, Lyapunov 函数 $F(\mathbf{\Omega}(t))$ 的变化量:

$$\begin{aligned} \Delta F(\mathbf{\Omega}(t)) &= F(\mathbf{\Omega}(t+1)) - F(\mathbf{\Omega}(t)) = \\ &= -\nabla E^T(\mathbf{\Omega}(t)) \Delta \mathbf{\Omega}(t) + \\ &\quad \frac{1}{2} \Delta \mathbf{\Omega}^T(t) \nabla^2 E(\mathbf{\Omega}(t)) \Delta \mathbf{\Omega}(t) \end{aligned} \quad (32)$$

$$\nabla E(\mathbf{\Omega}(t)) = \mathbf{\Theta}(\mathbf{\Omega}(t)) \quad (33)$$

$$\nabla^2 E(\mathbf{\Omega}(t)) = \mathbf{Q}(\mathbf{\Omega}(t)) + \lambda(t) \mathbf{I} \quad (34)$$

式中: $\nabla E(\Delta \mathbf{\Omega}(t))$ 和 $\nabla^2 E(\Delta \mathbf{\Omega}(t))$ 分别为误差函数的一阶导数、二阶导数矩阵。根据式(17)的更新规

则,将得到

$$\Delta \mathbf{\Omega}(t) = (\mathbf{Q}(\mathbf{\Omega}(t)) + \lambda(t) \mathbf{I})^{-1} \mathbf{\Theta}(t) \quad (35)$$

现在有以下收敛定理:

假设 RC-RBF 神经网络中的隐含层神经元数为固定 J ,同时网络参数根据式(17)中的规则进行更新,如果满足以下假设:

$$\|\Delta \mathbf{\Omega}(t)\| \leq \min \left\{ \|\Delta \mathbf{\Omega}(t-1)\|, \frac{\mathbf{\Theta}(\mathbf{\Omega}(t-1))}{\mathbf{Q}(\mathbf{\Omega}(t-1))} \right\} \quad (36)$$

那么,结合等式(32)~(34)可以得到

$$\Delta F(\mathbf{\Omega}(t)) = -\frac{1}{2} \Delta \mathbf{\Omega}^T(t) \nabla^2 E(\mathbf{\Omega}(t)) \Delta \mathbf{\Omega}(t) \quad (37)$$

$\nabla^2 E(\Delta \mathbf{\Omega}(t))$ 是正定的,当满足假设条件,于是得出:

$$\Delta F(\mathbf{\Omega}(t)) < 0 \quad (38)$$

由此可以得出 Lyapunov 函数 $F(\mathbf{\Omega}(t))$ 不是增加的,进一步得出当 $e(t) \rightarrow 0$,网络收敛

$$\lim_{t \rightarrow \infty} e(t) = 0 \quad (39)$$

综上所述,通过网路误差补偿更新神经元的参数,对神经元增长和删减两阶段的收敛性证明;同时也对结构固定阶段的 RBF 网络的收敛性也进行证明,因此提出设计方法的收敛性得以验证。

4 仿真实验

RC-RBF 神经网络能够根据研究对象的复杂动态变化在线调整隐含层神经元的个数,提高网络的预测能力,为验证算法的有效性和可行性,对 Mackey-class 时间序列和污水处理关键出水参数氨氮预测进行预测实验,对其算法进行验证。

利用均方根误差函数作为衡量网路的性能指标函数,计算公式如式(40)所示。

$$\text{RMSE}(t) = \sqrt{\frac{1}{2P} \sum_{p=1}^P (y_p(t) - o_p(t))^2} \quad (40)$$

式中: P 为样本总数, $y_p(t)$ 为第 p 个样本 t 时刻对应的网络输出, $o_p(t)$ 为第 p 个样本 t 时刻对应的期望输出。

1) Mackey-Glass 时间序列预测

Mackey-Glass 时间序列预测是一个典型的验证自组织网络性能的基准函数^[18]。其微分方程表达式如式(41)所示:

$$x(t+1) = (1-a)x(t) + \frac{bx(t-\tau)}{1+x^{10}(t-\tau)} \quad (41)$$

式中: $a=0.1$, $b=0.2$, $\tau=17$, 并且初始条件为 $x(0)=1.2$, $p=6$, $\Delta t=6$ 。过去的4个值 $\{x(t), x(t-\Delta t), x(t-2\Delta t), x(t-3\Delta t)\}$ 去预测 $x(t+p)$ 的值,预测模型如式(42)所示:

$$x(t+p) = f[x(t), x(t-\Delta t), x(t-2\Delta t), x(t-3\Delta t)] \quad (42)$$

根据式(42)产生1000个数据,其中, $t \in [136, 535]$ 产生500组作为训练数据, $t \in [636, 1135]$ 产

生的 500 组用作测试数据。网络的初始结构为 4-3-1, 阈值 ε 为 0.1。图 2~5 的仿真结果分别为训练过程中隐节点数目的变化, 网络测试输出和实际输出曲线, Mackey-class 测试误差曲线, 训练 RMSE 变化曲线。

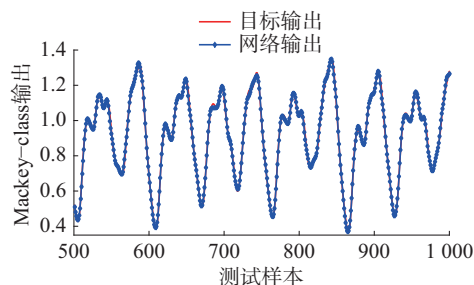


图 2 Mackey-class 测试输出和实际输出

Fig. 2 The curve of Mackey-class test output and actual output

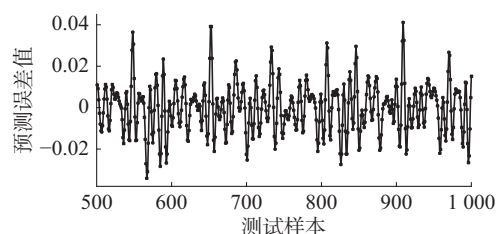


图 3 Mackey-class 测试误差曲线

Fig. 3 The curve of Mackey-class testing error

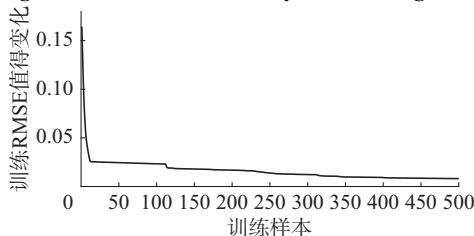


图 4 训练 RMSE 值的变化曲线

Fig. 4 The curve of training RMSE value

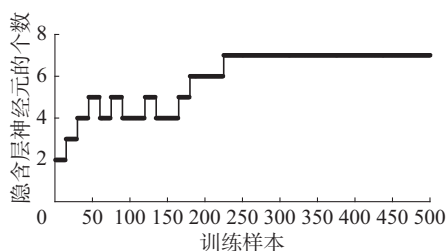


图 5 训练过程中隐节点的数目变化

Fig. 5 The number of hidden layer numbers

从图 2 可以看出, 网络的测试输出与实际输出曲线基本吻合, 说明该方法能够较好地逼近非线性时间序列。图 3 中显示网络的测试误差曲线, 误差值范围保持在 $[-0.04, 0.04]$ 较小的范围内, 表现出良好的泛化性能。图 4 显示的网络训练 RMSE 变化曲线, 误差随着训练过程一直保持下降趋势, 证明表 2 算法能够快速收敛到满意的性能。

为了验证算法的有效性, 该方法与其他自组

织神经网络进行对比, NFN-FOESA^[19], 自组织 RBF 神经网络 (self-organizing RBF neural networks, SORBF)^[20], 增长-修剪 RBF 神经网络 (growing and pruning RBF neural networks, GAP-RBF)^[21]、RBF-AFSII^[22]、SOFMLS^[23]。表 1 展现了不同算法的性能指标, 从表中可以看出, 和其他方法相比, 该方法能够采用较精简的结构很好地逼近非线性时间序列。

表 1 不同算法的性能对比

Table 1 The performance of different algorithms

算法	隐节点数目	训练RMSE	测试RMSE
RC-RBF	7	0.008 5	0.008 6
NFN-FOESA ^[19]	10*	0.118 3	0.094 6
SORBF ^[20]	12	0.01	0.010 0
GAP-RBF ^[21]	19	0.01	0.030 0
RBF-AFSII ^[22]	21	0.010 7	0.012 8
SOFMLS ^[23]	7	—	0.047 1

表 2 不同算法的性能对比

Table 2 The performance of different algorithms

算法	预测准确度/%	训练RMSE	测试RMSE
RC-RBF	98.84	0.023 7	0.030 6
ILM-RBF ^[24]	94.55	0.147 5	0.187 0
RBF ^[24]	90.63	0.287 0	0.317 3
BP ^[24]	88.42	0.384 1	0.475 1

2) 污水处理关键水质参数氨氮预测

出水水质参数预测模型的建立对于污水处理中减少微生物风险和过程的优化控制提供了一定的管理决策支持依据。由于污水处理过程中影响硝化和反硝化的因素众多, 动力学反应及其复杂, 导致影响氨氮的参数很多, 而且各参数之间相互作用, 具有强烈的耦合和非线性特性, 因此很难精确建立其机理模型; 而且, 氨氮的检测, 现在大多采用实验室取样离线分析, 操作繁琐, 需要很长的时间才能检测出等原因, 不利于实时准确获得水质参数。因此, 采用数据驱动的方式, 利用上述方法根据污水处理厂实际输入输出数据对氨氮这一出水参数建立准确快速的软测量模型。

实验过程中, 数据取自北京市某污水处理厂 2014 年的真实测量数据, 剔除异常数据并进行数据归一化之后, 140 组数据作为训练数据, 50 组数据作为测试数据。为消除输入数据间的高度相关性通过主元分析的方法, 选出对氨氮影响较大的输入变量, 依次为温度 T (Temperature), 好氧前段 DO (Dissolved Oxygen), 好氧末端总固体悬浮物 TSS (Total suspended solids concentration), 出水 PH

(Acidity and basicity), 厌氧末端 ORP (Oxidation-reduction potential), 泥龄。输出变量为出水氨氮。网络的初始结构为 6-2-1, 删减阈值设为 0.1。预测性能通过预测精度 pa (prediction accuracy) 来衡量。

$$pa = \frac{\sum_{t=1}^N (1 - |y_p(t) - o(t)|/o(t))}{N} \quad (43)$$

预测精度 pa 越大, 说明基于 RC-RBF 神经网络的氨氮预测模型性能越好。实验仿真结果如图 6~10。

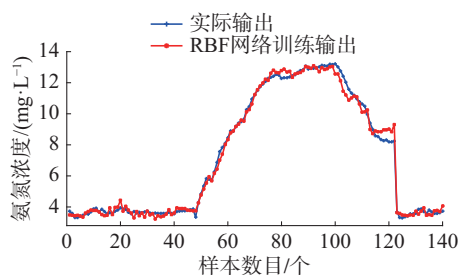


图 6 网络训练输出和实际输出曲线

Fig. 6 Network training output and Actual output

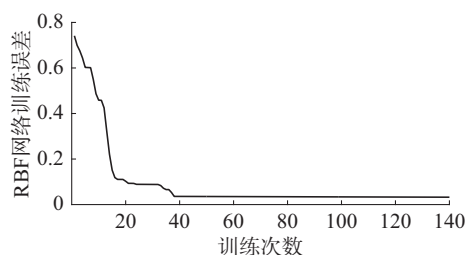


图 7 训练 RMSE 值变化

Fig. 7 The curve of training RMSE value

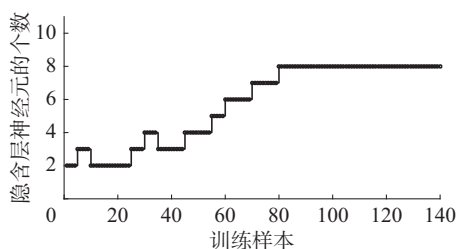


图 8 训练过程中隐节点的数目变化

Fig. 8 The number of hidden layer number

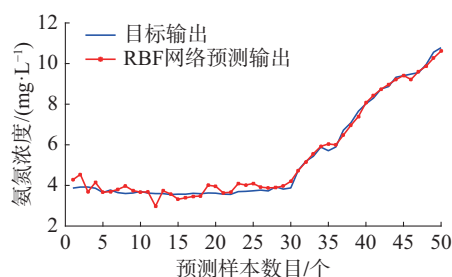


图 9 网络测试输出和实际输出

Fig. 9 The curve of test the actual output

从仿真结果可以看出, 图 6 显示的 RBF 网络训练输出和氨氮实际输出曲线, 曲线基本重合,

可以看出算法对污水处理动态过程具有较好的学习能力。图 7 表明算法在训练过程中 RMSE 的值呈现下降趋势, 训练时间为仅为 1.24 s, 快速收敛到较小的期望误差。图 9 显示 RBF 网络测试输出和实际输出变化曲线, 测试输出值和氨氮实际值基本吻合。从图 10 网络测试误差曲线看出, 测试误差较小, 说明对于实际非线性动态变化过程, RC-RBF 神经网络表现出良好的泛化能力。为了说明该方法针对氨氮预测的有效性, 还和其他方法进行对比, 采用固定神经元的改进 LM 算法 (Improved LM algorithm, ILM) 优化的 RBF 网络 (ILM-RBF), RBF 神经网络, BP 网络 3 个对比实验。表 2 显示了算法的性能对比。从中可以看出, 和其他方法相比, 该方法对于氨氮的预测精度更高, 训练和测试误差都较小。

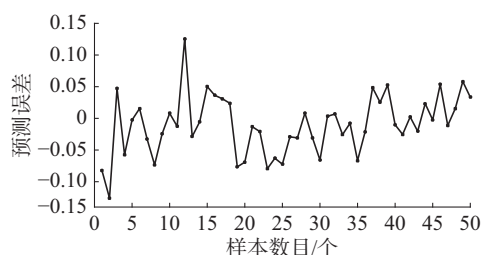


图 10 网络测试误差曲线

Fig. 10 The curve of network testing error

基于以上仿真结果, 可以得出基于改进的自适应 LM 算法的自组织 RBF 神经网络能够很好地对污水处理氨氮指标快速准确地在线预测, 并且具有较小的测试误差和较高的预测精度, 对于污水处理出水参数的实时监测和控制具有重要的意义, 有利于污水处理厂的正常运行。

5 结束语

文中通过采用相对贡献指标来设计网络结构; 同时, 采用改进的 LM 算法训练 RBF 神经网络的参数, 使得网络能够根据非线性系统准确快速地在在线结构调整和参数优化, 并表现出良好的泛化能力和预测精度, 一定程度上提高了 RBF 神经网络的非线性系统的建模性能和算法的鲁棒性。

1) 提出基于相对贡献指标结合误差信息处理能力来最大程度的挖掘隐含层神经元对输出神经元的相对贡献, 作为判断网络结构是否删减的条件, 并且对增加和删减阶段的神经元参数进行更新补偿; 同时, 证明网络结构删减和增加阶段, 神经元固定 3 个阶段的收敛性, 证明结构动态变化过程是收敛的。

2) 采用改进的 LM 算法, 将矩阵运算转化为向量相乘的方式, 避免整个雅可比矩阵的存储,

有效地减少了训练时间和矩阵的存储空间,自适应学习率的引入加快算法的收敛速度,并且提高了算法的预测性能。

3) 采用改进 LM 算法的 RC-RBF 神经网络通过对 Mackey-class 时间序列预测和污水处理过程的出水参数氨氮进行建模预测,仿真实验可以看出网络对于解决实际问题非线性函数逼近实验都具有较好的泛化能力和预测精度。

综上所述,本文提出的相对贡献指标结合网络信息处理能力能够有效解决 RBF 网络的结构调整问题;同时采用具有快速收敛特性的改进 LM 算法对网络参数优化,非线性时间序列仿真实验和污水处理氨氮预测实验都证明算法具有较好的预测精度,对于实际非线性系统的建模与应用有一定的理论意义和可行性。

参考文献:

- [1] QIAO J F, ZHANG Z Z, BO Y C. An online self-adaptive modular neural network for time-varying systems[J]. *Neurocomputing*, 2014, 125: 7–16.
- [2] 韩红桂, 李森, 乔俊飞. 基于模型输出敏感度分析的动态神经网络结构设计[J]. *电子学报*, 2010, 38(3): 731–736.
- [3] HAN Honggui, LI Miao, QIAO Junfei. Design of dynamic neural network based on the sensitivity analysis of model output[J]. *Acta electronica sinica*, 2010, 38(3): 731–736.
- [4] PLATT J. A resource-allocating network for function interpolation[J]. *Neural computation*, 1991, 3(2): 213–225.
- [5] LU Y G, SUNDARARAJAN N, SARATCHANDRAN P. A sequential learning scheme for function approximation using minimal radial basis function neural networks[J]. *Neural computation*, 1997, 9(2): 461–478.
- [6] PANCHAPAKESAN C, PALANISWAMI M, RALPH D, et al. Effects of moving the center's in an RBF network[J]. *IEEE transactions on neural networks*, 2002, 13(6): 1299–1307.
- [7] HUANG G B, SARATCHANDRAN P, SUNDARARAJAN N. A generalized growing and pruning RBF (GGAP-RBF) neural network for function approximation[J]. *IEEE transactions on neural networks*, 2005, 16(1): 57–67.
- [8] GONZALEZ J, ROJAS I, ORTEGA J, et al. Multi-objective evolutionary optimization of the size, shape, and position parameters of radial basis function networks for function approximation[J]. *IEEE transactions on neural networks*, 2003, 14(6): 1478–1495.
- [9] FENG H M. Self-generation RBFNs using evolutionary PSO learning[J]. *Neurocomputing*, 2006, 70(1): 241–251.
- [10] HAO C, YU G, XIA H. Online modeling with tunable RBF network[J]. *IEEE transactions on cybernetics*, 2013, 43(3): 935–947.
- [11] LIAN J M, LEE Y G, SCOTT D, SUDHOFF, et al. Self-organizing radial basis function network for real-time approximation of continuous-time dynamical systems[J]. *IEEE transactions on neural networks*, 2008, 19(3): 460–474.
- [12] YU H, REINER P D, XIE T T, et al. An incremental design of radial basis function networks[J]. *IEEE transactions on neural networks and learning systems*, 2014, 25(10): 1793–1803.
- [13] CONSTANTINOPOULOU C, LIKAS A. An incremental training method for the probabilistic RBF network[J]. *IEEE transactions on neural networks*, 2006, 17(4): 966–974.
- [14] CHEN S, HANZO L, TAN S. Symmetric complex-valued RBF receiver for multiple-antenna-aided wireless systems[J]. *IEEE transactions on neural networks*, 2008, 19(9): 1659–1665.
- [15] 王晓丽, 黄蕾, 杨鹏, 等. 动态 RBF 神经网络在浮选过程模型失配中的应用[J]. *化工学报*, 2016, 3(67): 897–902.
- [16] WANG Xiaoli, HUANG Lei, YANG Peng, et al. Dynamic RBF neural networks for model mismatch problem and its application in flotation process[J]. *CIESC journal*, 2016, 3(67): 897–902.
- [17] WILAMOWSKI B M, YU H. Improved computation for Levenberg–Marquardt training[J]. *IEEE transactions on neural networks*, 2010, 21(6): 930–937.
- [18] MA C, JIANG L. Some research on Levenberg–Marquardt method for the nonlinear equations[J]. *Applied mathematics and computation*, 2007, 184(2): 1032–1040.
- [19] XIE T T, YU H, HEWLETT J, et al. Fast and efficient second-order method for training radial basis function networks[J]. *IEEE transactions on neural networks and learning systems*, 2012, 23(4): 609–619.
- [20] QIAO J F, HAN H G. Identification and modeling of nonlinear dynamical systems using a novel self-organizing RBF-based approach[J]. *Automatica*, 2012, 48(8): 1729–1734.
- [21] CHEN C, WANG F Y. A self-organizing neuro-fuzzy network based on first order effect sensitivity analysis[J]. *Neurocomputing*, 2013, 118: 21–32.
- [22] HAN H G, CHEN Q L, QIAO J F. An efficient self-organizing RBF neural network for water quality prediction[J]. *Neural networks the official journal of the international neural network society*, 2011, 24(7): 717–725.
- [23] HUANG G B, SARATCHANDRAN P, SUNDARARAJAN N. An efficient sequential learning algorithm for growing and pruning RBF (GAP-RBF) networks[J]. *IEEE transactions on systems, man, and cybernetics, Part B: (Cybernetics)*, 2004, 34(6): 2284–2292.
- [24] CHO K B, WANG B Y. Radial basis function based adaptive fuzzy systems and their applications to system identification and prediction[J]. *Fuzzy sets and systems*, 1996, 83(3): 325–339.
- [25] EBADZADEH M M, SALIMI-BADR A. CFNN: correlation

ated fuzzy neural network[J]. Neurocomputing, 2015, 148: 430-444.

- [24] An R, Li W J, Han H G, et al. An improved Levenberg-Marquardt algorithm with adaptive learning rate for RBF neural network[C]//IEEE Control Conference, [s.l.], 2016: 3630-3635.

作者简介:



乔俊飞,男,1968年生,教授,博士生导师。主要研究方向为智能信息处理、智能优化控制。获教育部科技进步奖一等奖和北京市科学技术三等奖各1项,获得授权国家发明专利12项。发表学术论文近70篇,被SCI检索15篇。



安茹,女,1990年生,硕士研究生,主要研究方向为智能控制理论及应用、非线性系统建模。



韩红桂,男,1983年生,教授,博士生导师,主要研究方向为污水处理过程建模、优化与控制。获得授权国家发明专利13项,授权实用新型专利3项,授权软件著作权11项;研究成果获教育部科学技术进步奖一等奖等。近5年来发表学术论文35篇,撰写专著1部。

书讯:《探秘机器人王国》

由蔡自兴教授和翁环高级讲师合著的《探秘机器人王国》已由清华大学出版社隆重出版,向全国发行,正在京东和当当等书店热售中。

本书是一本以机器人学和人工智能知识和机器人技术为中心内容的科普及科幻长篇小说。书中通过形象与连续的故事和插图,介绍机器人的发展历史、基本结构与分类,在工矿业与农林业、空间与海洋探索、国防与安保、医疗卫生、家庭服务、文化娱乐、教育教学等方面的应用,以及智能化工厂、未来宇宙开发与星际航行和发展方向等。此外,还展望了其它一些新技术或潜在高新技术的未来应用。

本书密切联系实际,适当加入一些科学预测知识,故事情节生动,图文并茂,寓知识性、趣味性和娱乐性于一体,是广大青少年、大中小学生、中小学教师、机器人和人工智能产业园科技与工作人员以及从事科技与产业管理的政府与企业人员的课外阅读佳作,也是对机器人感兴趣的公众值得一看的好作品。小学高年级学生可以在家长和老师指导下阅读。

本书作者之一蔡自兴教授是一位在智能科技领域硕果累累、德高望重、著作等身和桃李满天下的科学家、教授。他在百忙中挤出宝贵时间,撰写科普小说和科普文章,普及机器人学和人工智能知识,培养广大青少年对科学技术的兴趣。他与夫人翁环老师合作编著这部《探秘机器人王国》科普与科幻著作长篇小说,是他们对科普教育的一个新贡献。读者通过阅读本书,能够了解机器人的过去、现在和将来,增进对机器人技术的兴趣与认识,因而能够对广大青少年读者有所裨益。

中国工程院院士、哈尔滨工业大学蔡鹤皋教授为本书作序,对本书给予高度评价。他说:本书的出版必将为广大读者,特别是青少年学生提供一份不可多得的精神快餐,为机器人学和人工智能的知识传播与普及发挥不可替代的重要作用,进而为我国建设智能强国贡献重要力量。

世界这么广大,想要探访的领域很多很多。厉害啦,机器人来了!让我们跟随《探秘机器人王国》的足迹去揭开机器人王国的神秘面纱,探索机器人家族的奥秘,拥抱多彩多姿的机器人,看看他们都干了哪些“雷倒众生”的新鲜事儿。