

## 基于随机配置网络的井下供给风量建模

王前进<sup>1</sup> 杨春雨<sup>2</sup> 马小平<sup>2</sup> 张春富<sup>1</sup> 彭思敏<sup>1</sup>

**摘要** 主通风机切换过程中, 取压风量测量作为监测井下供给风量的主要手段, 是矿井主扇通风系统安全、稳定与经济运行的重要保障。然而, 由于取压孔极易出现堵塞现象, 需要频繁维护, 导致无法实时测量井下供给风量, 难以实现主通风机切换过程的闭环优化控制。同时, 随着隐含层节点数的增加, 基于随机配置网络 (Stochastic configuration network, SCN) 的估计模型存在过拟合和泛化能力差的缺点。为了解决上述问题, 结合正则化 (Regularization, R) 技术, 本文提出一种新型的改进 SCN 算法, 即 RSC 算法, 用于井下供给风量的建模。基准回归分析和工业实验表明: 与 SCN 方法相比, 建立的 RSC 模型具有较高的模型精度和较好的泛化性能。

**关键词** 主通风机, 切换过程, 井下供给风量, 随机配置网络, 正则化

**引用格式** 王前进, 杨春雨, 马小平, 张春富, 彭思敏. 基于随机配置网络的井下供给风量建模. 自动化学报, 2021, 47(8): 1963–1975

**DOI** 10.16383/j.aas.c190602

## Underground Airflow Quantity Modeling Based on SCN

WANG Qian-Jin<sup>1</sup> YANG Chun-Yu<sup>2</sup> MA Xiao-Ping<sup>2</sup> ZHANG Chun-Fu<sup>1</sup> PENG Si-Min<sup>1</sup>

**Abstract** In a main fan switchover process (MFSP), airflow quantity measuring by pressure drop across a duct is the main means for monitoring the distribution of underground airflow quantity (UAQ). It ensures the safe, stable and economic operation of the mine main fan ventilation system. However, as the pressure port is prone to blocking and requires frequent maintenance, it is impossible to realize the real-time measurement of UAQ and closed-loop optimal control of MFSP. Meanwhile, with the increase of number of hidden nodes, the estimation model based on stochastic configuration network (SCN) has the disadvantages of overfitting and poor generalization ability. To solve the above problems, this paper develops a novel improved SCN algorithm by incorporating the regularization (R) technique called RSC algorithm. Benchmark regression analysis and industrial tests show that compared with SCN, the established RSC model possesses higher model accuracy and better generalization ability.

**Key words** Main fan, switchover process, underground airflow quantity, stochastic configuration network, regularization

**Citation** Wang Qian-Jin, Yang Chun-Yu, Ma Xiao-Ping, Zhang Chun-Fu, Peng Si-Min. Underground airflow quantity modeling based on SCN. *Acta Automatica Sinica*, 2021, 47(8): 1963–1975

主通风机切换过程 (Main fan switchover process, MFSP) 被广泛用来保证矿井的持续安全生产<sup>[1]</sup>。矿井必须配有成对的主通风机, 采用轮流运行的工

作模式。其中, 运行的一台叫工作主通风机, 另一台叫备用主通风机。井下供给风量作为主通风机切换过程的关键运行指标, 对井下作业影响很大, 因此有必要对其进行实时监测, 保证主通风机切换过程的平稳运行以提供充足的井下供给风量。然而, 由于恶劣的井下工作环境, 取压孔容易出现堵塞现象, 导致无法实时准确地监测井下供给风量状态。另外, 堵塞的取压孔必须等到大周期 (一般半年或一年一次) 的主通风机检修才能进行取压孔的清理。因此有必要建立井下供给风量的准确可靠性模型, 为矿井值守人员提供井下工作状况和供给风量信息。

目前井下供给风量模型主要分为两大类: 机理模型和数据驱动模型。主通风机切换过程涉及到两大动力源的过渡, 涉及到 4 个风门的切换, 涉及到井下供给风流的控制等。因此, 井下供给风量机理模型包括: 主通风机模型可用其特性曲线来表示<sup>[2]</sup>;

收稿日期 2019-08-24 录用日期 2019-11-16

Manuscript received August 24, 2019; accepted November 16, 2019

国家自然科学基金 (61873272, 61603392), 江苏省自然科学基金 (BK20191043), 江苏省“双创团队”项目 (2017), 盐城工学院校级科研项目 (xjr2019018) 资助

Supported by National Natural Science Foundation of China (61873272, 61603392), Natural Science Foundation of Jiangsu Province (BK20191043), Jiangsu Dual Creative Teams Programme Project (2017), Funding for School-Level Research Projects of Yancheng Institute of Technology (xjr2019018)

本文责任编辑 刘艳军

Recommended by Associate Editor LIU Yan-Jun

1. 盐城工学院电气工程学院 盐城 224051 2. 中国矿业大学信息与控制工程学院 徐州 221116

1. School of Electrical Engineering, Yancheng Institute of Technology, Yancheng 224051 2. School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116

对于风门支路,首先通过数据拟合方法建立风门风阻与风门开度之间的非线性关系,再采用定常<sup>[3-4]</sup>或非定常<sup>[5-6]</sup>不可压缩 Navier-Stokes 方程来建模;对于地下矿井,由于复杂多变的井下工作环境和错综复杂的网络布局以及频繁变化的设备状态,难以建立模型.通常首先将其近似为一个长度确定的风流支路,再采用类似于风门支路的建模方法来建模.井下供给风量模型是由上述主要结构的子模型组成,通过基尔霍夫风压和风流定律来建立.然而,由于机理模型一般是基于一定的理论假设而建立的,如地下矿井的近似化处理,实际主通风机切换过程要远远复杂于机理模型,因此机理模型在实际主通风机切换过程中往往难以应用.数据驱动的井下供给风量建模方法,不需要掌握切换过程的复杂变化,仅利用数据就可以建立井下供给风量的估计模型,因此数据驱动井下供给风量建模是目前主通风机切换过程建模研究的热点.如,文献[7]利用静态工况下的井下供给风量和风门分支风门风阻的数据,建立了基于 RBF 神经网络的井下供给风量模型.

近年,由于具有极快的学习速度,且能有效地解决传统神经网络收敛速度慢、易陷入局部极小的问题,增量式随机权网络 (Incremental random weight network, IRWN) 已被广泛用于复杂工业过程的建模<sup>[8-10]</sup>. IRWN 本质上是一种采用构造算法得到的单隐层前馈神经网络 (Single layer feedforward neural network, SLFNN), 比传统神经网络具有更简单的结构和更高的计算效率<sup>[11]</sup>. 针对 IRWN 收敛性问题,文献[12]通过对一类 IRWN 进行理论分析和仿真研究,得出一个结论: IRWN 的万能逼近能力只有在一定的约束条件下才能得到保证.因此,寻找保证 IRWN 收敛性的约束条件是建立 IRWN 的关键.如,文献[13]提出了一种基于不等式约束条件的随机配置网络 (Stochastic configuration network, SCN), 其不仅具有万能逼近能力,而且具有较快的学习速度.与标准 IRWN 相比, SCN 算法具有较快的学习速度和较高的模型精度,但是 SCN 算法依然存在如下的问题:

**问题.** 随着隐含层节点数目的增加, SCN 算法可能存在过拟合和泛化性能差的问题.由于隐含层节点的增加,使得模型学习时结构较复杂,致使模型对已知训练数据表现较好,而对未知数据具有较差的测试效果,造成过拟合的问题.由于过拟合的存在,使得模型鲁棒性不足、泛化能力差,难以实际应用.

另外, PLC 控制系统已广泛应用于主通风机切换过程<sup>[14]</sup>. 由于受 PLC 存储空间、运行效率等的限

制,需要建立一个先进的小尺寸网络来减轻计算和存储压力.因此有必要建立高质量、小尺寸的井下供给风量估计模型.

针对上述问题,在 SCN 算法的基础上,结合正则化 (Regularization, R) 技术,本文提出一种新型的改进 SCN 算法,即 RSC 算法.该算法包含两个版本: RSC-I 和 RSC-II.两个版本的算法在相同的不等式约束条件下随机分配隐性参数,但采用不同的方式计算输出权重. RSC-I 算法通过求解局部正则化最小二乘问题来确定输出权重,而 RSC-II 算法采用全局正则化最小二乘法计算输出权重.基准回归分析和工业实验表明:相比 SCN 算法,本文方法不仅具有较高的建模精度和较好的泛化能力,而且还减轻了 PLC 存储和计算的壓力.

## 1 主通风机切换过程与特性分析

### 1.1 主通风机切换过程及井下供给风量建模的必要性

主通风机切换过程如图 1 所示,由两大动力源的过渡、4 个风门的切换 2 部分组成.假设当前工作主通风机为一号主通风机,备用主通风机为二号主通风机,那么一号主通风机、完全关闭的一号水平风门与二号垂直风门、完全打开的一号垂直风门和地下矿井共同构成一个典型的矿井通风系统.主通风机切换过程就是切断一号主通风机与地下矿井的连接,同时建立二号主通风机与地下矿井的连接.矿井通常采用的主通风机切换过程可描述如下:首先,启动二号主通风机,再打开二号垂直风门和一号水平风门;同时,关闭二号水平风门和一号垂直风门;最后,直到 4 个风门完全到位后,才关闭一号主通风机.

主通风机切换时,井下供给风量是由两台主通风机共同提供的,同时又受井下工作环境的影响,对实现主通风机切换过程闭环优化控制起着重要作用.井下供给风量不仅是表征能耗水平和运行状况的重要指标,而且决定地下矿井工作环境质量、采煤效率和矿井安全生产.目前,国内外对主通风机切换过程闭环控制和运行优化的研究较少<sup>[15-17]</sup>,并且这些研究成果严重依赖井下供给风量的实时测量.但是在实际应用中,由于井下供给风流含有成分复杂的气、液、固三相混合杂质,致使取压孔容易发生堵塞,造成风量测量装置的测量效果不理想.因此,对井下供给风量进行有效估计,不仅能为矿井值守人员提供信息来判断主通风机切换过程运行状态和井下工作状况,而且有益于实现井下供给风

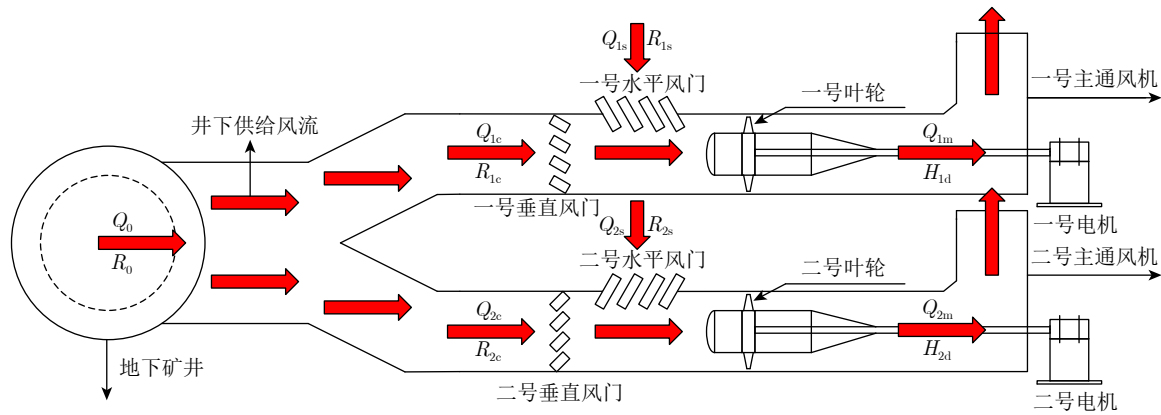


图 1 主通风机切换过程示意图

Fig.1 Diagram of a main fan switchover process

量的控制和运行优化.

## 1.2 建模复杂性及特性分析

主通风机切换过程是一个极其复杂的过程. 其中, 包含气、液、固三相混合杂质的井下供给风流耦合交错, 同时井下工作环境极其恶劣; 地下矿井局部通风机的启停和井下控制气流的风门状态以及作业人员的变动, 运行工况不稳定; 各类采矿活动 (回采工作面、电池和燃料充电站、矿石破碎站等) 中, 污染物产生的速度及其质量的变化, 运行工况动态时变. 这些导致了难以建立准确可靠的主通风机切换过程模型.

表 1 列出了主通风机切换过程的相关变量. 从图 1 可知, 影响井下供给风量指标的关键过程变量为:  $H_{1d}$ 、 $H_{2d}$ 、 $R_{1c}$ 、 $R_{1s}$ 、 $R_{2c}$ 、 $R_{2s}$ 、 $R_0$ . 下面对上述 7 个变量进行具体地分析.

1) 一号主通风机压头  $H_{1d}$  和二号主通风机压头  $H_{2d}$  对井下供给风量  $Q_0$  的影响

对于主通风机, 在额定转速下, 其风压  $H_f$  可描述为主通风机风量  $Q_f$  的函数. 根据该主通风机特性曲线<sup>[18]</sup>, 一般将这个函数描述为如下二次多项式:

$$H_f = a_2 Q_f^2 + a_1 Q_f + a_0 \quad (1)$$

式中,  $a_0$ ,  $a_1$  和  $a_2$  为待定常数, 根据风机特性曲线进行确定.

根据风机定律<sup>[19]</sup>, 转速  $N$  所对应的主通风机风压  $H$  可以写成:

$$H = a_2 Q^2 + a_1 Q \left( \frac{N}{N_f} \right) + a_0 \left( \frac{N}{N_f} \right)^2 \quad (2)$$

式中,  $N_f$  为额定转速,  $Q$  为转速  $N$  所对应的主通风机风量.

从式 (2) 可知, 当转速已知时, 通过主通风机风

表 1 主通风机切换过程相关变量  
Table 1 Related variables in the MFSP

变量	符号	单位
井下供给风量	$Q_0$	$\text{m}^3/\text{s}$
井下风阻	$R_0$	$\text{kg}/\text{m}^7$
一号垂直风门风量	$Q_{1c}$	$\text{m}^3/\text{s}$
一号垂直风门风阻	$R_{1c}$	$\text{kg}/\text{m}^7$
一号水平风门风量	$Q_{1s}$	$\text{m}^3/\text{s}$
一号水平风门风阻	$R_{1s}$	$\text{kg}/\text{m}^7$
一号主通风机风量	$Q_{1m}$	$\text{m}^3/\text{s}$
一号主通风机压头	$H_{1d}$	Pa
二号垂直风门风量	$Q_{2c}$	$\text{m}^3/\text{s}$
二号垂直风门风阻	$R_{2c}$	$\text{kg}/\text{m}^7$
二号水平风门风量	$Q_{2s}$	$\text{m}^3/\text{s}$
二号水平风门风阻	$R_{2s}$	$\text{kg}/\text{m}^7$
二号主通风机风量	$Q_{2m}$	$\text{m}^3/\text{s}$
二号主通风机压头	$H_{2d}$	Pa

压可以得到对应的主通风机风量. 因此, 对于转速可测的两台主通风机, 可得:

$$Q_{1m} = f_1(H_{1d}), Q_{2m} = f_2(H_{2d}) \quad (3)$$

式中,  $f_1(\cdot)$  和  $f_2(\cdot)$  为确定的非线性函数.

根据基尔霍夫风流定律<sup>[20]</sup>, 从图 1 可得:

$$Q_{1m} = Q_{1c} + Q_{1s} \quad (4)$$

$$Q_{2m} = Q_{2c} + Q_{2s} \quad (5)$$

$$Q_0 = Q_{1c} + Q_{2c} \quad (6)$$

由式 (3), (4) 和式 (5) 可知,  $H_{1d}$  的变化会引起  $Q_{1c}$  的变化; 同时当  $H_{2d}$  发生变化时,  $Q_{2c}$  也将随之发生改变. 另外, 由式 (6) 可知,  $Q_0$  是由  $Q_{1c}$  和  $Q_{2c}$  共同决定的. 因此,  $H_{1d}$  和  $H_{2d}$  对井下供给风量  $Q_0$  产生重要影响.

2) 风门支路风门风阻  $R_{1c}$ ,  $R_{1s}$ ,  $R_{2c}$  和  $R_{2s}$  对井下供给风量  $Q_0$  的影响

由 Darcy-Weisbach 方程可知, 风道压降与通过风道的风量之间的函数关系是非线性的<sup>[2]</sup>. 根据文献 [21-22], 可将这个函数关系写成:

$$H_{ic} = R_{ic}Q_{ic}^2, H_{is} = R_{is}Q_{is}^2, i = 1, 2 \quad (7)$$

式中,  $H_{ic}$  和  $H_{is}$  分别为主通风机垂直风门和水平风门压降,  $i$  为主通风机.

由于  $H_{ic}$  和  $H_{is}$  可看作  $i$  号主通风机的负载, 因此它们是由  $H_{id}$  提供的. 另外, 由式 (7) 可知, 风量和风阻之间呈现非线性关系. 那么上式可写成:

$$\begin{aligned} Q_{1s} &= \eta(R_{1s}, H_{1d}), Q_{1c} = \eta(R_{1c}, H_{1d}) \\ Q_{2s} &= \eta(R_{2s}, H_{2d}), Q_{2c} = \eta(R_{2c}, H_{2d}) \end{aligned} \quad (8)$$

式中,  $\eta(\cdot)$  为非线性函数.

由式 (4), (5) 和式 (8) 可知, 风门风阻  $R_{1s}$ ,  $R_{1c}$ ,  $R_{2c}$  和  $R_{2s}$  是影响风量  $Q_{1s}$ ,  $Q_{1c}$ ,  $Q_{2c}$  和  $Q_{2s}$  的重要因素. 再结合式 (6) 可知, 上述风阻是影响井下供给风量  $Q_0$  的关键过程变量.

3) 井下风阻  $R_0$  对井下供给风量  $Q_0$  的影响

对于地下矿井, 压降  $H_0$  与风阻  $R_0$  之间呈现出明显的非线性特征. 用公式描述为:

$$H_0 = R_0Q_0^2 \quad (9)$$

$H_0$  可视为两台主通风机的负载, 其数值由  $H_{1d}$  和  $H_{2d}$  共同决定. 主通风机切换过程中,  $R_0$  是不需要进行调节的, 但会受到一些因素的影响. 例如, 井下工作人员的变动, 井下控制气流的风门状态, 井下作业状况等. 上述这些因素都会引起  $R_0$  的变化. 根据式 (9) 可知,  $R_0$  的变化会引起井下供给风量  $Q_0$  的改变.

综上所述, 将主通风机压头  $H_{1d}$  和  $H_{2d}$ 、风门风阻  $R_{1s}$ ,  $R_{1c}$ ,  $R_{2c}$  和  $R_{2s}$ 、井下风阻  $R_0$  作为建模输入变量, 就可以实现井下供给风量  $Q_0$  的估计. 这个非线性映射可描述为:

$$Q_0 = \varphi(H_{1d}, H_{2d}, R_{1s}, R_{1c}, R_{2c}, R_{2s}, R_0) \quad (10)$$

式中,  $\varphi(\cdot)$  为未知的非线性函数.

## 2 相关算法

### 2.1 随机配置网络

SCN 开始于一个小尺寸的 SLFNN, 首先在一个不等式约束条件下, 逐个选取隐含层节点, 同时自适应地选择隐性参数的范围; 然后, 利用全局最小二乘法确定当前网络的输出权重; 最后, 直到达到设定的最大隐含层节点数目或期望的网络精度时

结束隐含层节点的增加.

令  $\Gamma := \{g_1, g_2, g_3, \dots\}$  表示一组实值函数,  $\text{span}(\Gamma)$  表示  $\Gamma$  张成的函数空间; 令  $L_2(D)$  表示定义在  $D \subset \mathbf{R}^d$  上的所有勒贝格可测函数  $f = [f_1, f_2, \dots, f_m]: \mathbf{R}^d \rightarrow \mathbf{R}^m$  的空间.

给定一个目标函数  $f: \mathbf{R}^d \rightarrow \mathbf{R}^m$ , 假设存在一个建立好的带有  $L-1$  个隐含层节点的 SLFNN, 那么 SLFNN 的输出可表示为:

$$f_{L-1}(x) = \sum_{j=1}^{L-1} \beta_j g_j(w_j^T x + b_j), f_0 = 0 \quad (11)$$

式中,  $L = 1, 2, \dots$ ,  $w_j$  和  $b_j$  分别为第  $j$  个隐含层节点的输入权值和偏置,  $\beta_j = [\beta_{j1}, \beta_{j2}, \dots, \beta_{jm}]^T$  为连接第  $j$  个隐含层节点与输出层节点的输出权值,  $g_j$  为第  $j$  个隐含层节点的激活函数.

根据式 (11), 可以得到当前网络残差  $e_{L-1}$ :

$$e_{L-1} = f - f_{L-1} = [e_{L-1,1}, e_{L-1,2}, \dots, e_{L-1,m}] \quad (12)$$

SCN 算法构造一个不等式约束条件来增加第  $L$  个隐含层节点, 其不等式约束的形式如下:

$$\langle e_{L-1,q}, g_L \rangle^2 \geq b_g^2 \delta_{L,q}, q = 1, 2, \dots, m \quad (13)$$

式中,  $\forall g \in \Gamma$ ,  $b_g \in \mathbf{R}^+$ , 使得  $0 < \|g\| < b_g$  成立,  $\delta_L$  的形式为:

$$\delta_L = \sum_{q=1}^m \delta_{L,q}, \delta_{L,q} = (1 - r - u_L) \|e_{L-1,q}\|^2 \quad (14)$$

式中,  $0 < r < 1$ ,  $u_L$  满足  $\lim_{L \rightarrow +\infty} u_L = 0$  和  $0 < u_L \leq (1 - r)$ .

SCN 中表现最好的 SCN-III 算法采用全局最小二乘法来计算输出权重:

$$[\beta_1^*, \beta_2^*, \dots, \beta_L^*] = \arg \min \|f - \sum_{j=1}^L \beta_j g_j\| \quad (15)$$

SCN 中表现最好的 SCN-III 算法采用上述方式进行隐含层节点的逐个增加, 直到满足设定的条件.

### 2.2 正则化随机权神经网络

与标准随机神经网络 (Random vector functional-link networks, RVFLNs) 相比, 正则化 RVFLNs 不仅能有效避免过拟合问题, 而且能通过减小输出权重来降低模型复杂性.

对于给定的神经网络训练样本  $(x_i, t_i)_{i=1}^N$ , 正则化 RVFLNs 的目标代价函数为:

$$\begin{aligned} \min : J &= \frac{1}{2} \|\beta\|_2^2 + \frac{C}{2} \sum_{i=1}^N \epsilon_i^2 \\ \text{s.t.} : h(\mathbf{x}_i)\beta &= \mathbf{t}_i - \epsilon_i, \forall i \end{aligned} \quad (16)$$

式中,  $\epsilon_i$  为网络残差,  $C$  为正则化系数.

通过正则化最小二乘法直接计算输出权重  $\beta$ , 具体结果如下:

$$\beta = \begin{cases} \left( \mathbf{H}^T \mathbf{H} + \frac{\mathbf{I}}{C} \right)^{-1} \mathbf{H}^T \mathbf{Y}, & N \geq L \\ \mathbf{H}^T \left( \mathbf{H}^T \mathbf{H} + \frac{\mathbf{I}}{C} \right)^{-1} \mathbf{Y}, & N < L \end{cases} \quad (17)$$

式中,  $\mathbf{H} = [\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_L]$ ,  $\mathbf{Y} = [\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_N]^T$ .

### 3 RSC 算法

#### 3.1 RSC 算法的收敛性分析和证明

给定神经网络训练样本  $(\mathbf{x}, \mathbf{t})$ , 其中,  $\mathbf{x} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$ ,  $\mathbf{x}_i = [x_{i,1}, x_{i,2}, \dots, x_{i,d}]^T \in \mathbf{R}^d$ ,  $\mathbf{t} = \{\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_N\}$ ,  $\mathbf{t}_i = [t_{i,1}, t_{i,2}, \dots, t_{i,m}]^T \in \mathbf{R}^m$ , 假设已经搭建一个带有  $L-1$  个隐含层节点的 SLFNN, 那么可以得到当前网络的输出权重:

$$\beta = [\beta_1, \beta_2, \dots, \beta_{L-1}] \quad (18)$$

接下来, 需要构造一个方案来分配隐性参数, 并计算输出权重  $\beta_L$ , 使得建立的带有  $L$  个隐含层节点的 SLFNN 具有更高的模型精度. 因此, 引入如下目标代价函数:

$$\begin{aligned} f(\beta_L) &= \frac{1}{2} [\beta \beta_L] \begin{bmatrix} \beta \\ \beta_L \end{bmatrix} + \frac{C}{2} \|e_L\|^2 = \\ &= \frac{1}{2} \|\beta\|_2^2 + \frac{1}{2} \|\beta_L\|_2^2 + \frac{C}{2} \|e_L\|^2 \end{aligned} \quad (19)$$

将式 (19) 两边同时对  $\beta_L$  求导, 再结合  $e_L = e_{L-1} - \beta_L \mathbf{g}_L$ , 可得:

$$\begin{aligned} \frac{\partial f(\beta_L)}{\partial \beta_L} &= \beta_L - C(e_{L-1} - \beta_L \mathbf{g}_L) \mathbf{g}_L = \\ &= \beta_L - C \langle e_{L-1}, \mathbf{g}_L \rangle + C \beta_L \mathbf{g}_L^2 \end{aligned} \quad (20)$$

令式 (20) 等于 0, 可得:

$$\beta_L = \frac{\langle e_{L-1}, \mathbf{g}_L \rangle}{\|\mathbf{g}_L\|^2 + \frac{1}{C}} \quad (21)$$

**定理 1.** 假设  $\text{span}(\mathbf{\Gamma})$  在  $L_2$  空间中稠密, 且  $\forall \mathbf{g} \in \mathbf{\Gamma}$ ,  $0 < \|\mathbf{g}\| < b_g$ , 其中,  $b_g \in \mathbf{R}^+$ . 对于一个给

定的正实数  $C$ , 记  $\varepsilon = \frac{\|b_g\|^2}{\left(\|b_g\|^2 + \frac{1}{C}\right)^2 / \left(\|b_g\|^2 + \frac{2}{C}\right)}$ .

给定  $1 - \varepsilon < r < 1$ , 一个非负实数序列  $u_L$  满足  $1 -$

$\varepsilon - r \leq u_L \leq (1 - r)$ . 对于  $L = 1, 2, \dots$  和  $q = 1, 2, \dots, m$ , 做出如下定义:

$$\delta_L = \sum_{q=1}^m \delta_{L,q}, \delta_{L,q} = (1 - r - u_L) \|\mathbf{e}_{L-1,q}\|^2 \quad (22)$$

$$\xi = \lim_{L \rightarrow \infty} (1 - \varepsilon)^L < 1 \quad (23)$$

$$\gamma = \lim_{L \rightarrow \infty} \prod_{k=1}^L (r + u_k) < 1 \quad (24)$$

若生成的随机基函数  $\mathbf{g}_L$  使得下式成立:

$$\langle \mathbf{e}_{L-1,q}, \mathbf{g}_L \rangle^2 \geq \left(b_g^2 + \frac{1}{C}\right)^2 / \left(b_g^2 + \frac{2}{C}\right) \delta_{L,q} \quad (25)$$

输出权重  $\beta_{L,q}$  由如下公式获取:

$$\beta_{L,q} = \frac{\langle \mathbf{e}_{L-1,q}, \mathbf{g}_L \rangle}{\|\mathbf{g}_L\|^2 + \frac{1}{C}} \quad (26)$$

那么,  $\sqrt{\xi} \|\mathbf{f}\| \leq \lim_{L \rightarrow +\infty} \|\mathbf{f} - \mathbf{f}_L\| \leq \sqrt{\gamma} \|\mathbf{f}\|$ , 其中,  $\mathbf{f}_L = \sum_{j=1}^L \beta_j \mathbf{g}_j$ ,  $\beta_j = [\beta_{j,1}, \beta_{j,2}, \dots, \beta_{j,m}]^T$ .

**证明.** 由式 (26) 可得:

$$\begin{aligned} \|\mathbf{e}_L\|^2 - \|\mathbf{e}_{L-1}\|^2 &= \\ &= \sum_{q=1}^m (\langle \mathbf{e}_{L-1,q} - \beta_{L,q} \mathbf{g}_L, \mathbf{e}_{L-1,q} - \beta_{L,q} \mathbf{g}_L \rangle - \\ &= \langle \mathbf{e}_{L-1,q}, \mathbf{e}_{L-1,q} \rangle) = \\ &= \sum_{q=1}^m (-2 \langle \mathbf{e}_{L-1,q}, \beta_{L,q} \mathbf{g}_L \rangle + \langle \beta_{L,q} \mathbf{g}_L, \beta_{L,q} \mathbf{g}_L \rangle) = \\ &= - \sum_{q=1}^m (\langle \beta_{L,q} \mathbf{g}_L, \beta_{L,q} \mathbf{g}_L \rangle + \frac{2}{C} \langle \beta_{L,q}, \beta_{L,q} \rangle) \end{aligned} \quad (27)$$

整理式 (27) 得到:

$$\begin{aligned} \|\mathbf{e}_{L-1}\|^2 - \|\mathbf{e}_L\|^2 &= \\ &= \frac{\sum_{q=1}^m \langle \mathbf{e}_{L-1,q}, \mathbf{g}_L \rangle^2}{\left(\|\mathbf{g}_L\|^2 + \frac{1}{C}\right)^2 / \left(\|\mathbf{g}_L\|^2 + \frac{2}{C}\right)} \leq \\ &= \frac{\|\mathbf{e}_{L-1}\|^2 \|\mathbf{g}_L\|^2}{\left(\|\mathbf{g}_L\|^2 + \frac{1}{C}\right)^2 / \left(\|\mathbf{g}_L\|^2 + \frac{2}{C}\right)} \leq \\ &= \frac{\|\mathbf{e}_{L-1}\|^2 \|b_g\|^2}{\left(\|b_g\|^2 + \frac{1}{C}\right)^2 / \left(\|b_g\|^2 + \frac{2}{C}\right)} = \\ &= \varepsilon \|\mathbf{e}_{L-1}\|^2, 0 < \varepsilon < 1 \end{aligned} \quad (28)$$

上式可写成:

$$(1 - \varepsilon)\|e_{L-1}\|^2 \leq \|e_L\|^2 \tag{29}$$

根据式 (22), (25) 和式 (26) 可得:

$$\begin{aligned} \|e_L\|^2 - (r + u_L)\|e_{L-1}\|^2 = & \sum_{q=1}^m (\langle e_{L-1,q} - \beta_{L,q}g_L, e_{L-1,q} - \beta_{L,q}g_L \rangle - \\ & (r + u_L)\langle e_{L-1,q}, e_{L-1,q} \rangle) = \\ & \sum_{q=1}^m ((1 - r - u_L)\langle e_{L-1,q}, e_{L-1,q} \rangle - \\ & 2\langle e_{L-1,q}, \beta_{L,q}g_L \rangle + \langle \beta_{L,q}g_L, \beta_{L,q}g_L \rangle) = \\ & (1 - r - u_L)\|e_{L-1}\|^2 + \sum_{q=1}^m \left( \frac{-2\langle e_{L-1,q}, g_L \rangle^2}{\|g_L\|^2 + \frac{1}{C}} + \right. \\ & \left. \frac{\langle e_{L-1,q}, g_L \rangle^2 \|g_L\|^2}{\left(\|g_L\|^2 + \frac{1}{C}\right)^2} \right) \end{aligned} \tag{30}$$

式 (30) 可写成:

$$\begin{aligned} \|e_L\|^2 - (r + u_L)\|e_{L-1}\|^2 = & (1 - r - u_L)\|e_{L-1}\|^2 - \\ & \frac{\sum_{q=1}^m \langle e_{L-1,q}, g_L \rangle^2}{\left(\|g_L\|^2 + \frac{1}{C}\right)^2 / \left(\|g_L\|^2 + \frac{2}{C}\right)} \leq \\ \delta_L - \frac{\sum_{q=1}^m \langle e_{L-1,q}, g_L \rangle^2}{\left(b_g^2 + \frac{1}{C}\right)^2 / \left(b_g^2 + \frac{2}{C}\right)} \leq 0 \end{aligned} \tag{31}$$

由式 (31) 可得:

$$\|e_L\|^2 \leq (r + u_L)\|e_{L-1}\|^2 \tag{32}$$

结合式 (29) 和式 (32) 可得:

$$(1 - \varepsilon)^L \|e_0\|^2 \leq \|e_L\|^2 \leq \prod_{k=1}^L (r + u_k) \|e_0\|^2 \tag{33}$$

由式 (23), (24) 和式 (33) 可得  $\xi \|f\|^2 \leq \lim_{L \rightarrow \infty} \|e_L\|^2 \leq \gamma \|f\|^2$ . 因此可得  $\sqrt{\xi} \|f\| \leq \lim_{L \rightarrow \infty} \|e_L\| \leq \sqrt{\gamma} \|f\|$ .  $\square$

**注 1.** 从式 (19) 可知, 正则化系数  $C$  起到平衡  $\|\beta_L\|_2^2$  和  $\|e_L\|_2^2$  的作用, 在保证模型精度的同时, 降低了复杂度, 因此需要对其进行适当选择. 本文通过简单的一维搜索程序来确定合适的正则化系数  $C$ . 从定理 1 可知, 对于给定合适的  $C$ , RSC 算法不仅

保留了 SCN 算法收敛速度快的优点, 而且能有效地克服过拟合问题.

从定理 1 可知,  $\beta_L = [\beta_{L,1}, \beta_{L,2}, \dots, \beta_{L,m}]^T$  是通过计算  $\beta_{L,q} = \langle e_{L-1,q}, g_L \rangle / (\|g_L\|^2 + \frac{1}{C})$  而得到的, 并且在增加下一个隐含层节点时保持不变. 但是, 这种构造方法会导致 RSC 算法收敛速度变慢. 因此, 本文给出输出权重的重新计算方法. 具体过程是首先根据式 (25) 确定  $g_j$  ( $j = 1, 2, \dots, L$ ), 再采用全局正则化最小二乘法计算输出权重  $\beta_1, \beta_2, \dots, \beta_L$ . 令  $[\beta_1^*, \beta_2^*, \dots, \beta_L^*] = \arg \min_{\beta} \frac{C}{2} \left\| f - \sum_{j=1}^L \beta_j g_j \right\|^2 + \frac{1}{2} \|\beta\|_2^2$   $e_L^* = f - \sum_{j=1}^L \beta_j^* g_j = [e_{L,1}^*, e_{L,2}^*, \dots, e_{L,m}^*]$ . 定义中间变量  $\tilde{\beta}_{L,q} = \langle e_{L-1,q}^*, g_L \rangle / \left(\|g_L\|^2 + \frac{1}{C}\right)$ ,  $\tilde{e}_L = e_{L-1}^* - \tilde{\beta}_L g_L$ , 其中,  $q = 1, 2, \dots, m$ ,  $\tilde{\beta}_L = [\tilde{\beta}_{L,1}, \tilde{\beta}_{L,2}, \dots, \tilde{\beta}_{L,m}]^T$ ,  $e_0^* = f$ .

**定理 2.** 假设  $\text{span}(\Gamma)$  在  $L_2$  空间中稠密, 且  $\forall g \in \Gamma, 0 < \|g\| < b_g$ , 其中,  $b_g \in \mathbf{R}^+$ . 对于一个给定的正实数  $C$ , 存在一个足够小的正实数  $\varepsilon(C)$ . 给定  $1 - \varepsilon < r < 1$ , 一个非负实数序列  $u_L$  满足  $1 - \varepsilon - r \leq u_L \leq (1 - r)$ . 对于  $L = 1, 2, \dots$  和  $q = 1, 2, \dots, m$ , 做出如下定义:

$$\delta_L^* = \sum_{q=1}^m \delta_{L,q}^*, \delta_{L,q}^* = (1 - r - u_L) \|e_{L-1,q}^*\|^2 \tag{34}$$

$$\xi = \lim_{L \rightarrow \infty} (1 - \varepsilon)^L < 1 \tag{35}$$

$$\gamma = \lim_{L \rightarrow \infty} \prod_{k=1}^L (r + u_k) < 1 \tag{36}$$

若生成的随机基函数  $g_L$  满足下式:

$$\langle e_{L-1,q}^*, g_L \rangle^2 \geq \frac{\left(b_g^2 + \frac{1}{C}\right)^2}{\left(b_g^2 + \frac{2}{C}\right)} \delta_{L,q}^* \tag{37}$$

输出权重  $\beta_{L,q}$  由如下公式计算:

$$[\beta_1^*, \beta_2^*, \dots, \beta_L^*] = \arg \min_{\beta} \frac{C}{2} \left\| f - \sum_{j=1}^L \beta_j g_j \right\|^2 + \frac{1}{2} \|\beta\|_2^2 \tag{38}$$

那么,  $\sqrt{\xi} \|f\| \leq \lim_{L \rightarrow +\infty} \|f - f_L^*\| \leq \sqrt{\gamma} \|f\|$ , 其中,  $f_L^* = \sum_{j=1}^L \beta_j^* g_j$ ,  $\beta_j^* = [\beta_{j,1}^*, \beta_{j,2}^*, \dots, \beta_{j,m}^*]^T$ .

**证明.** 容易可知,  $\|\mathbf{e}_L^*\|^2 \leq \|\tilde{\mathbf{e}}_L\|^2 = \|\mathbf{e}_{L-1}^* - \|\mathbf{e}_{L-1}^*\|^2 \leq \|\tilde{\mathbf{e}}_{L-1}\|^2$ . 那么可以得到:

$$\begin{aligned} & \|\mathbf{e}_L^*\|^2 - \|\mathbf{e}_{L-1}^*\|^2 \leq \\ & \|\tilde{\mathbf{e}}_L\|^2 - \|\mathbf{e}_{L-1}^*\|^2 = \\ & \sum_{q=1}^m (\langle \mathbf{e}_{L-1,q}^* - \tilde{\beta}_{L,q} \mathbf{g}_L, \mathbf{e}_{L-1,q}^* - \tilde{\beta}_{L,q} \mathbf{g}_L \rangle - \\ & \langle \mathbf{e}_{L-1,q}^*, \mathbf{e}_{L-1,q}^* \rangle) = \\ & \sum_{q=1}^m (-2\langle \mathbf{e}_{L-1,q}^*, \tilde{\beta}_{L,q} \mathbf{g}_L \rangle + \langle \tilde{\beta}_{L,q} \mathbf{g}_L, \tilde{\beta}_{L,q} \mathbf{g}_L \rangle) = \\ & \sum_{q=1}^m \left( \frac{-2\langle \mathbf{e}_{L-1,q}^*, \mathbf{g}_L \rangle^2}{\|\mathbf{g}_L\|^2 + \frac{1}{C}} + \frac{\langle \mathbf{e}_{L-1,q}^*, \mathbf{g}_L \rangle^2 \|\mathbf{g}_L\|^2}{\left(\|\mathbf{g}_L\|^2 + \frac{1}{C}\right)^2} \right) = \\ & - \frac{\sum_{q=1}^m \langle \mathbf{e}_{L-1,q}^*, \mathbf{g}_L \rangle^2}{\left(\|\mathbf{g}_L\|^2 + \frac{1}{C}\right)^2 / \left(\|\mathbf{g}_L\|^2 + \frac{2}{C}\right)} \quad (39) \end{aligned}$$

由于  $L_2$  范数会产生有偏估计, 存在一个充分小的正实数  $\theta$ , 使得  $\|\mathbf{e}_L^*\|^2 > \theta$  成立. 那么可得:

$$\begin{aligned} & \|\mathbf{e}_{L-1}^*\|^2 - \|\mathbf{e}_L^*\|^2 \leq \\ & \|\mathbf{e}_{L-1}^*\|^2 - \theta \quad (40) \end{aligned}$$

令  $\|\mathbf{e}_{L-1}^*\|^2 - \theta = \varepsilon \|\mathbf{e}_{L-1}^*\|^2$ ,  $0 < \varepsilon < 1$  通过式 (40) 可得:

$$(1 - \varepsilon) \|\mathbf{e}_{L-1}^*\|^2 \leq \|\mathbf{e}_L^*\|^2 \quad (41)$$

由式 (34), (37) 和 (39) 可得:

$$\begin{aligned} & \|\mathbf{e}_L^*\|^2 - (r + u_L) \|\mathbf{e}_{L-1}^*\|^2 \leq \\ & \|\tilde{\mathbf{e}}_L\|^2 - (r + u_L) \|\mathbf{e}_{L-1}^*\|^2 = \\ & \sum_{q=1}^m (\langle \mathbf{e}_{L-1,q}^* - \tilde{\beta}_{L,q} \mathbf{g}_L, \mathbf{e}_{L-1,q}^* - \tilde{\beta}_{L,q} \mathbf{g}_L \rangle - \\ & (r + u_L) \langle \mathbf{e}_{L-1,q}^*, \mathbf{e}_{L-1,q}^* \rangle) = \\ & \sum_{q=1}^m ((1 - r - u_L) \langle \mathbf{e}_{L-1,q}^*, \mathbf{e}_{L-1,q}^* \rangle - \\ & 2\langle \mathbf{e}_{L-1,q}^*, \tilde{\beta}_{L,q} \mathbf{g}_L \rangle + \langle \tilde{\beta}_{L,q} \mathbf{g}_L, \tilde{\beta}_{L,q} \mathbf{g}_L \rangle) = \\ & (1 - r - u_L) \|\mathbf{e}_{L-1}^*\|^2 + \sum_{q=1}^m \left( \frac{-2\langle \mathbf{e}_{L-1,q}^*, \mathbf{g}_L \rangle^2}{\|\mathbf{g}_L\|^2 + \frac{1}{C}} + \right. \end{aligned}$$

$$\begin{aligned} & \left. \frac{\langle \mathbf{e}_{L-1,q}^*, \mathbf{g}_L \rangle^2 \|\mathbf{g}_L\|^2}{\left(\|\mathbf{g}_L\|^2 + \frac{1}{C}\right)^2} \right) = \\ & (1 - r - u_L) \|\mathbf{e}_{L-1}^*\|^2 - \\ & \frac{\sum_{q=1}^m \langle \mathbf{e}_{L-1,q}^*, \mathbf{g}_L \rangle^2}{\left(\|\mathbf{g}_L\|^2 + \frac{1}{C}\right)^2 / \left(\|\mathbf{g}_L\|^2 + \frac{2}{C}\right)} \leq \\ & \delta_L - \frac{\sum_{q=1}^m \langle \mathbf{e}_{L-1,q}^*, \mathbf{g}_L \rangle^2}{\left(b_g^2 + \frac{1}{C}\right)^2 / \left(b_g^2 + \frac{2}{C}\right)} \leq 0 \quad (42) \end{aligned}$$

类似于定理 1 的证明可得  $\sqrt{\xi} \|\mathbf{f}\| \leq \lim_{L \rightarrow \infty} \|\mathbf{e}_L^*\| \leq \sqrt{\gamma} \|\mathbf{f}\|$ .

**注 2.** 从式 (38) 可知, 正则化系数  $C$  控制  $\|\beta\|_2^2$  和  $\|\mathbf{e}_L\|_2^2$  之间的平衡. 另外, 不等式约束 (37) 的作用是增大网络残差函数  $\|\mathbf{e}_L\|_2^2$  的下降量, 从而加快算法的收敛速度. 因此, 定理 2 所对应的 RSC 算法不仅能有效抑制过拟合现象, 而且具有较快的收敛速度.

### 3.2 算法描述

定理 1 和定理 2 分别对应本文提出的 RSC 算法的两个版本, 即 RSC-I 算法和 RSC-II 算法. 由于 RSC-II 算法比 RSC-I 算法具有更快的收敛速度, 因此本文着重给出 RSC-II 算法的实现步骤.

令  $\mathbf{e}_{L-1}(\mathbf{x}) = [\mathbf{e}_{L-1,1}(\mathbf{x}), \mathbf{e}_{L-1,2}(\mathbf{x}), \dots, \mathbf{e}_{L-1,m}(\mathbf{x})] \in \mathbf{R}^{N \times m}$  为带有  $L-1$  个隐含层节点的 SLFNN 残差, 其中,  $\mathbf{e}_{L-1,q}(\mathbf{x}) = [e_{L-1,q}(\mathbf{x}_1), e_{L-1,q}(\mathbf{x}_2), \dots, e_{L-1,q}(\mathbf{x}_N)]^T \in \mathbf{R}^N$ ,  $q = 1, 2, \dots, m$ . 做出如下定义:

$$\begin{aligned} \mathbf{h}_L(\mathbf{x}) = & [\mathbf{g}_L(\mathbf{w}_L^T \mathbf{x}_1 + b_L), \mathbf{g}_L(\mathbf{w}_L^T \mathbf{x}_2 + b_L), \dots, \\ & \mathbf{g}_L(\mathbf{w}_L^T \mathbf{x}_N + b_L)]^T \quad (43) \end{aligned}$$

式中,  $\mathbf{h}_L(\mathbf{x})$  为第  $L$  个隐含层节点输出. 那么当前网络的隐含层输出矩阵可写为  $\mathbf{H}_L = [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_L]$ .

事实上,  $\beta_{L,q} = \langle \mathbf{e}_{L-1,q}, \mathbf{g}_L \rangle / \left(\|\mathbf{g}_L\|^2 + \frac{1}{C}\right)$  ( $q = 1, 2, \dots, m$ ) 可写成如下形式:

$$\beta_{L,q} = \frac{\mathbf{e}_{L-1,q}(\mathbf{x})^T \cdot \mathbf{h}_L(\mathbf{x})}{\mathbf{h}_L^T(\mathbf{x}) \cdot \mathbf{h}_L(\mathbf{x}) + \frac{1}{C}}, \quad q = 1, 2, \dots, m \quad (44)$$

为简单起见, 引入一组变量  $\xi_{L,q}$ ,  $q = 1, 2, \dots, m$ . 其表达式如下:

$$\xi_{L,q} = \frac{(e_{L-1,q}^T(\mathbf{x}) \cdot \mathbf{h}_L(\mathbf{x}))^2}{\Upsilon} - (1-r-u_L) \cdot e_{L-1,q}^T(\mathbf{x})e_{L-1,q}(\mathbf{x}) \quad (45)$$

式中,  $\Upsilon = \left( \mathbf{h}_L^T(\mathbf{x}) \cdot \mathbf{h}_L(\mathbf{x}) + \frac{1}{C} \right)^2 / \left( \mathbf{h}_L^T(\mathbf{x}) \times \mathbf{h}_L(\mathbf{x}) + \frac{2}{C} \right)$ .

RSC-II 算法流程:

给定神经网络的训练样本  $(\mathbf{x}, \mathbf{t})$ , 其中,  $\mathbf{x} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$ ,  $x_i \in \mathbf{R}^d$ ,  $\mathbf{t} = \{\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_N\}$ ,  $\mathbf{t}_i \in \mathbf{R}^m$ ; 最大隐含层节点数为  $L_{\max}$ , 期望精度为  $\epsilon$ , 随机配置的最大次数为  $T_{\max}$ ; 选择一组正的标量  $\Upsilon = \{\lambda_{\min} : \Delta\lambda : \lambda_{\max}\}$ .

**步骤 1.** 神经网络的初始化阶段: 令  $\mathbf{e}_0 := [\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_N]^T$ , 给定  $C > 0$ , 一个充分小的正实数  $\epsilon$  给定  $1 - \epsilon < r < 1$ , 两个空集:  $\Omega$  and  $\mathbf{W}$ ;

**步骤 2.** 神经网络的训练阶段: 随机参数配置和输出权重计算.

当  $L \leq L_{\max}$  和  $\|\mathbf{e}_0\|_F > \epsilon$  时,

随机参数配置:

1) **For**  $\lambda \in \Upsilon$ ,

2) **For**  $k = 1, 2, \dots, T_{\max}$ ,

3) 分别从  $[-\lambda, \lambda]^d$  和  $[-\lambda, \lambda]$  中随机选取

$\mathbf{w}_L$  和  $b_L$ ;

4) 分别由式 (43) 和式 (45) 得到  $\mathbf{h}_L$  和  $\xi_{L,q}$ , 并令  $u_L = (1-r)/(L+1)$ ;

5) **If**  $\min \xi_{L,1}, \xi_{L,2}, \dots, \xi_{L,m} \geq 0$

6) 将  $\mathbf{w}_L$  和  $b_L$  存储在  $\mathbf{W}$  中,  $\xi_L =$

$\sum_{q=1}^m \xi_{L,q}$  存储在  $\Omega$  中;

7) **Else** 返回 2)

8) **End If**

9) **End For**

10) **If**  $\mathbf{W}$  不为空集

11) 在  $\Omega$  中找到最大  $\xi_L$  所对应的  $\mathbf{w}_L^*$  和  $b_L^*$ , 并令  $\mathbf{H}_L = [\mathbf{h}_1^*, \mathbf{h}_2^*, \dots, \mathbf{h}_L^*]$ ;

12) **Break**(返回 16));

13) **Else** 在  $\tau \in (0, 1-r)$  中随机产生  $\tau$ , 并更新  $r := r + \tau$ , 再返回 2);

14) **End If**

15) **End For**

输出权重计算:

16) 计算  $\beta^* = [\beta_1^*, \beta_2^*, \dots, \beta_L^*]^T := \left( \mathbf{H}_L^T \mathbf{H}_L + \frac{\mathbf{I}}{C} \right)^{-1} \times \mathbf{H}_L^T \mathbf{t}$ ;

17) 计算  $\mathbf{e}_L = \mathbf{t} - \mathbf{H}_L \beta^*$ ;

18) 令  $\mathbf{e}_0 := \mathbf{e}_L$ ,  $L := L + 1$ ;

算法结束, 并得到最优的  $\beta_1^*, \beta_2^*, \dots, \beta_L^*$ ,  $\mathbf{w}^* = [\mathbf{w}_1^*, \mathbf{w}_2^*, \dots, \mathbf{w}_L^*]$  和  $b^* = [b_1^*, b_2^*, \dots, b_L^*]$ .

## 4 实验分析

为验证所提 RSC-II 算法的有效性, 与 SCN-III<sup>[13]</sup> 算法进行比较分析. 为保证实验结果的可靠性, 选择一个函数近似、四个基准数据集和一组主通风切换过程数据作为仿真对象. 所有仿真实验在 MATLAB 2016a 中进行, 使用 CPU 为 i5, 3.4 GHz, 内存为 8 GB RAM 的 PC 机. 选择 Sigmoid 函数  $g(\mathbf{x}) = 1/(1 + \exp(-\mathbf{x}))$  作为两种算法隐含层节点的激活函数. 将神经网络的输入输出数据归一化到区间  $[0, 1]$ .

在一定的噪声水平下对两种算法进行评价. 选择均方根误差 (Root mean square error, RMSE) 作为评价算法泛化能力的指标. 每组进行 20 次试验, 选择 20 次试验的平均值 (Mean) 和标准差 (Standard deviation, STD) 作为最终结果. 随机参数的最大配置次数为  $L_{\max} = 200$ , 网络精度取值为  $\epsilon = 0.01$ . 其他相关参数的设定将在具体的仿真实验中给出.

### 4.1 函数近似

考虑一个函数近似<sup>[29]</sup> 问题:

$$y = 0.2e^{-(10x-4)^2} + 0.5e^{-(80x-40)^2} + 0.3e^{-(80x-20)^2} \quad (46)$$

式中,  $x \in [0, 1]$ .

由  $\text{rand}(0, 1)$  随机产生 1 000 个数据作为训练样本, 由  $\text{linspace}(0, 1, 300)$  等间隔生成 300 个数据组成测试样本. 为验证所提 RSC-II 算法能有效克服过拟合问题, 需要对训练数据进行预处理, 即在训练样本中添加一定比例的异常值. 训练样本中异常值的比例为 5%, 其形式为:

$$y_{i, \text{Outlier}} = y_i + \text{rand}(0, 1) - 0.5 \quad (47)$$

式中,  $y_i \in [0, 1]$  为归一化训练样本的第  $i$  个输出,  $i$  从  $[1, 2, \dots, N]$  中随机选取,  $\text{rand}(0, 1)$  为  $(0, 1)$  之间的随机数.

由式 (37) 和式 (38) 可知, 正则化系数  $C$  严重影响算法的学习精度,  $C$  的选择可以提高精度, 也可以降低精度. 为显示  $C$  对网络性能的影响, 针对不同的  $C$ , 对 RSC-II 算法进行仿真.  $C$  在集合  $\{2^{-20}, 2^{-19}, \dots, 2^0, \dots, 2^{19}, 2^{20}\}$  中依次取值, 一共有 41 个数值. 同时设定  $\lambda_{\min} = 100$ ,  $\Delta\lambda = 1$ ,  $L_{\max} = 50$ . 从图 2 可以看出,  $C$  越小, RSC-II 算法的训练能力越差, 对测试能力亦是如此. 当  $C$  大于  $2^0$  时, 与



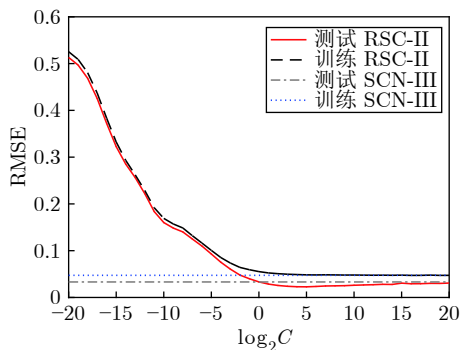


图 2 不同  $C$  下 RSC-II 和 SCN-III 算法对函数近似的训练与测试精度结果对比

Fig.2 The result comparison of training and testing accuracy of RSC-II and SCN-III on function approximation with varying  $C$

SCN-III 算法相比, RSC-II 算法显示出更好的测试性能. 另外, 对于一个合适的正则化系数  $C$ , 在相同的训练精度下, RSC-II 算法比 SCN-III 算法具有更高的测试精度. 因此, 在建立网络时, 需要确定合适的系数  $C$ . 那么在下面的仿真实验中, RSC-II 算法使用合适的正则化系数  $C$ .

图 3 (a) 为包含 5% 异常值的 1 000 个训练样本, 图 3 (b) 为两种算法对测试数据集的测试性能. 可以看出, 相比于 SCN-III 算法, RSC-II 算法具有更好的拟合能力. 从图 4 可知, RSC-II 算法总是获得比 SCN-III 算法更低的测试 RMSE. 随着隐含层节点数目的增加, RSC-II 的泛化能力得到提高或者基本保持不变, 而 SCN-III 的性能变差. 这是因为 SCN-III 算法始终存在过拟合问题, 并且随着隐含层节点的增加, 过拟合问题更加严重. RSC-II 算法通过引入正则化系数用于平衡近似精度和模型复杂度, 可有效避免过拟合问题. 表 2 给出了两种算法的具体实验结果.

### 4.2 基准数据

为进一步验证所提 RSC-II 算法的有效性, 选取 UCI 数据库<sup>[24]</sup> 中的四个数据集进行实验. 相关数据集的信息和参数设置如表 3 所示. 采用类似于函数近似问题中的策略进行训练数据预处理. 首先, 从归一化训练数据中随机选取 5% 的样本; 然后, 通过式 (47) 对选中的数据进行预处理.

从图 5 可以看出, 随着隐含层节点数的增加, SCN-III 算法的测试集 RMSE 呈现明显上升趋势, 而 RSC-II 算法的测试集 RMSE 趋于平稳, 因此 SCN-III 算法出现过拟合现象, 而 RSC-II 算法未出现过拟合问题. 这是因为随着隐含层节点数的增加, 神经网络结构变得过于复杂, 使得 SCN-III 算法出

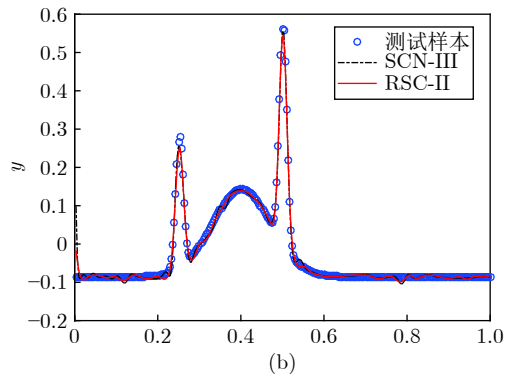
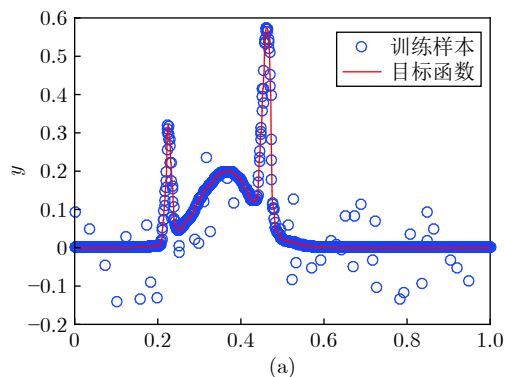


图 3 (a) 包含 5% 异常值的 1 000 个函数近似训练样本和目标函数; (b) 两种算法对测试数据的近似性能

Fig.3 (a) 1 000 training samples containing 5% outliers for function approximation and target function; (b) Approximation performance on the test dataset by two learning algorithms

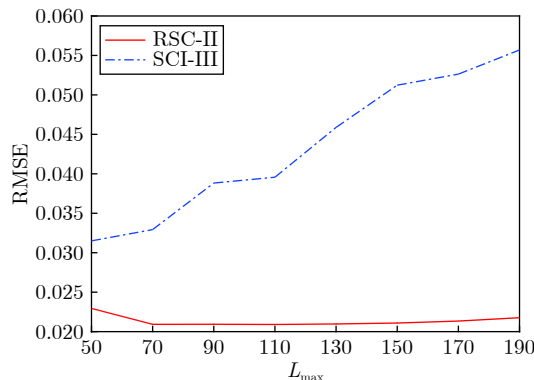


图 4 平均两种算法在测试数据集上的测试 RMSE  
Fig.4 Average testing RMSE of the two algorithms on the test dataset

现过拟合问题, 对测试数据表现较差. 由于 RSC-II 算法引入的正则化系数  $C$  起到了平衡模型复杂度与拟合精度的作用, 避免了过拟合问题, 对测试数据具有较好的测试效果. 两种算法的具体实验结果如表 4 所示, 其中, (a)、(b)、(c)、(d) 分别为 Wine

表 2 函数近似的性能比较  
Table 2 Performance comparisons on the function approximation

算法	不同 $L_{\max}$ 所对应的测试性能 (Mean, STD)							
	50	70	90	110	130	150	170	190
SCN-III	0.0315, 0.0032	0.0329, 0.0033	0.0388, 0.0048	0.0396, 0.0032	0.0459, 0.0043	0.0512, 0.0058	0.0526, 0.0048	0.0557, 0.0095
RSC-II	<b>0.0229, 0.0019</b>	<b>0.0209, 0.0013</b>	<b>0.0209, 0.0008</b>	<b>0.0209, 0.0003</b>	<b>0.0210, 0.0005</b>	<b>0.0211, 0.0004</b>	<b>0.0213, 0.0004</b>	<b>0.0218, 0.0003</b>

表 3 回归数据集与参数设置

Table 3 Specifications of benchmark problems and some parameter settings

数据集	属性	输出	训练数据	测试数据	$\lambda_{\min}$	$\Delta\lambda$
Wine Quality	12	1	3918	980	1	0.1
Concrete	9	1	772	258	1	0.1
Yacht	7	1	227	81	1	1
Airfoil Self-noise	6	1	1352	151	1	0.1

Quality、Concrete、Yacht、Airfoil Self-noise 数据集.

### 4.3 井下供给风量估计

前面两个小节的实验仿真验证了所提 RSC-II 算法的有效性, 本小节将其在实际主通风机切换过程进行井下供给风量估计. 在实际工业现场采集 1 500 个样本作为建模的输入输出数据, 其中选择 1 400 个样本作为训练数据, 其余 100 个样本作为测试数据. 采用前面两个小节中类似的方式对训练

数据进行预处理,  $\lambda_{\min}$  值设定为 1,  $\Delta\lambda$  也取值为 1.

从图 6 可以看出, RSC-II 算法的泛化能力始终优于 SCN-III 算法的泛化能力. 当隐含层节点数为 50 时, 两种算法的性能比较相近. 随着隐含层节点数的增加, RSC-II 算法测试集的 RMSE 始终保持在较低的水平, 而 SCN-III 算法测试集的 RMSE 上升趋势明显. 因此, 相对于 RSC-II 算法, SCN-III 算法出现了较为明显的过拟合现象. 其原因在于当神经网络结构变得尤为复杂时, SCN-III 缺乏应对的能力, 从而导致泛化能力差、鲁棒性不足的问题. 引入了正则化系数  $C$  的 RSC-II 算法对模型精度和复杂度进行折中, 在保证精度的同时, 降低了模型复杂度. 表 5 列出了两种算法实现的具体实验结果.

图 7 给出了在  $L_{\max} = 50$  时两种算法的测试性能. 可以看出, 相比于 SCN-III 算法, RSC-II 算法的估计值能很好地拟合实际值, 模型精度较高. 综上, 采用 RSC-II 方法建立的井下供给风量估计模型可以以很高的精度对建模数据进行拟合, 尤其是

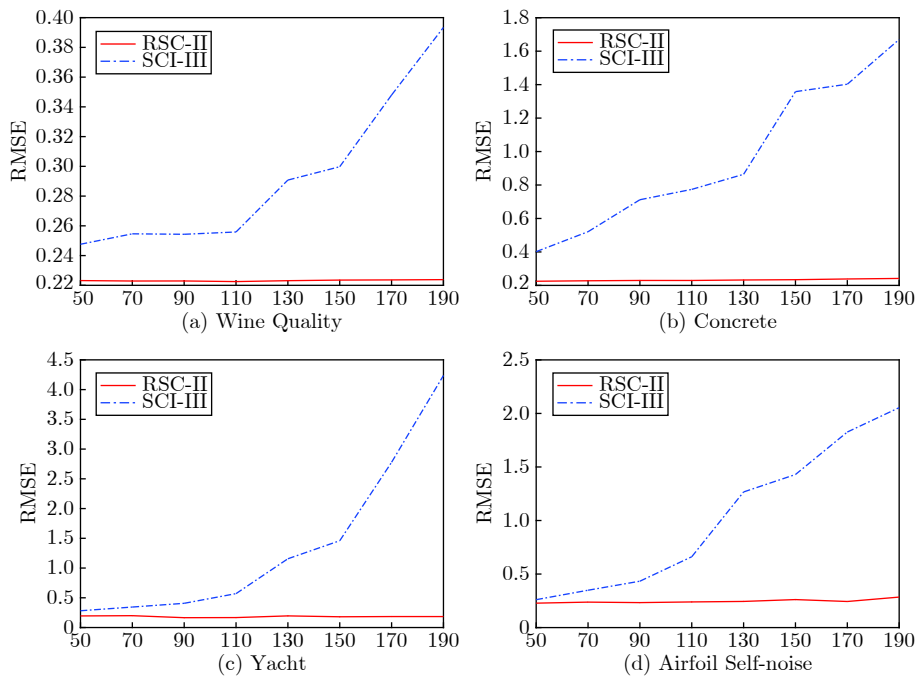


图 5 不同  $L_{\max}$  下 RSC-II 和 SCN-III 算法对基准数据集的测试 RMSE 对比

Fig. 5 Test RMSE comparison of the RSC-II and SCN-III algorithms on four benchmark datasets with different  $L_{\max}$

表 4 对基准数据集的性能对比  
Table 4 Performance comparisons on benchmark datasets

数据集	算法	不同 $L_{max}$ 所对应的测试性能 (Mean, STD)							
		50	70	90	110	130	150	170	190
(a)	SCN-III	0.2476, 0.0233	0.2546, 0.0531	0.2543, 0.0492	0.2559, 0.0464	0.2908, 0.0612	0.2997, 0.0841	0.3480, 0.1484	0.3934, 0.1886
	RSC-II	<b>0.2232, 0.0011</b>	<b>0.2229, 0.0014</b>	<b>0.2229, 0.0016</b>	<b>0.2225, 0.0016</b>	<b>0.2231, 0.0017</b>	<b>0.2235, 0.0022</b>	<b>0.2236, 0.0021</b>	<b>0.2238, 0.0021</b>
(b)	SCN-III	0.4017, 0.1060	0.5209, 0.1587	0.7120, 0.2247	0.7739, 0.2010	0.8640, 0.2324	1.3580, 0.7945	1.4025, 0.4526	1.6700, 0.5415
	RSC-II	<b>0.2248, 0.0068</b>	<b>0.2277, 0.0059</b>	<b>0.2295, 0.0094</b>	<b>0.2295, 0.0076</b>	<b>0.2323, 0.0089</b>	<b>0.2341, 0.0090</b>	<b>0.2384, 0.0063</b>	<b>0.2416, 0.0056</b>
(c)	SCN-III	0.3167, 0.2289	0.3528, 0.1436	0.3820, 0.1781	0.6620, 0.2900	0.9062, 0.5600	1.2226, 0.3571	2.7450, 1.1718	4.0242, 1.0377
	RSC-II	<b>0.1308, 0.0209</b>	<b>0.1216, 0.0107</b>	<b>0.1197, 0.0152</b>	<b>0.1168, 0.0111</b>	<b>0.1164, 0.0118</b>	<b>0.1113, 0.0117</b>	<b>0.1056, 0.0112</b>	<b>0.1066, 0.0113</b>
(d)	SCN-III	0.2607, 0.0313	0.3489, 0.1007	0.4332, 0.1219	0.6615, 0.2487	1.2668, 0.4213	1.4295, 0.5883	1.8264, 0.5194	2.0539, 0.9744
	RSC-II	<b>0.2281, 0.0091</b>	<b>0.2378, 0.0147</b>	<b>0.2332, 0.0153</b>	<b>0.2395, 0.0214</b>	<b>0.2443, 0.0279</b>	<b>0.2615, 0.0402</b>	<b>0.2437, 0.0234</b>	<b>0.2848, 0.0512</b>

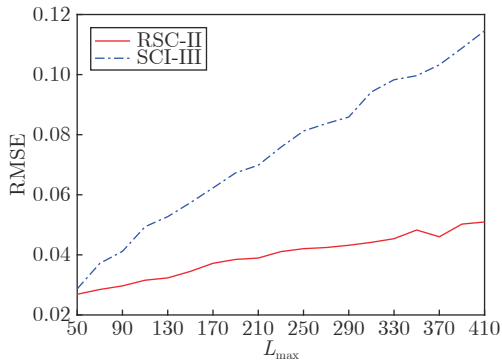


图 6 平均两种算法在实际 MFSP 数据集上的测试 RMSE  
Fig.6 Average testing RMSE of the two algorithms on the actual MFSP dataset

可以避免过拟合问题. 另外, 随着主通风机切换过程样本数据的增多, 需要适当地增加隐含层节点数. 结果表明, 基于 RSC-II 的建模算法在隐含层节点数为 170 时仍表现出良好的泛化性能和预测精度. 因此, 本文所提方法能够用于主通风机切换过程实现井下供给风量的建模.

### 5 结论

针对 SCN 算法建模时存在过拟合和泛化能力差的问题, 本文结合正则化技术, 提出一种新型的 RSC 算法, 用于建立主通风机切换过程井下供给风量的估计模型. 与 SCN 算法相比, 建立的 RSC-II 模型不仅具有更高的模型精度, 而且具有更好的泛化能力. 尤为重要的是 RSC-II 算法可以有效避免

表 5 在实际 MFSP 数据集上的性能对比  
Table 5 Performance comparisons on the actual MFSP dataset

算法	不同 $L_{max}$ 所对应的测试性能 (Mean, STD)							
	50	110	170	230	270	330	370	410
SCN-III	0.0287, 0.0029	0.0494, 0.0039	0.0623, 0.0069	0.0758, 0.0063	0.0837, 0.0059	0.0983, 0.0107	0.1033, 0.0096	0.1146, 0.0095
RSC-II	<b>0.0269, 0.0021</b>	<b>0.0316, 0.0017</b>	<b>0.0372, 0.0035</b>	<b>0.0411, 0.0046</b>	<b>0.0424, 0.0047</b>	<b>0.0454, 0.0066</b>	<b>0.0460, 0.0054</b>	<b>0.0509, 0.0077</b>

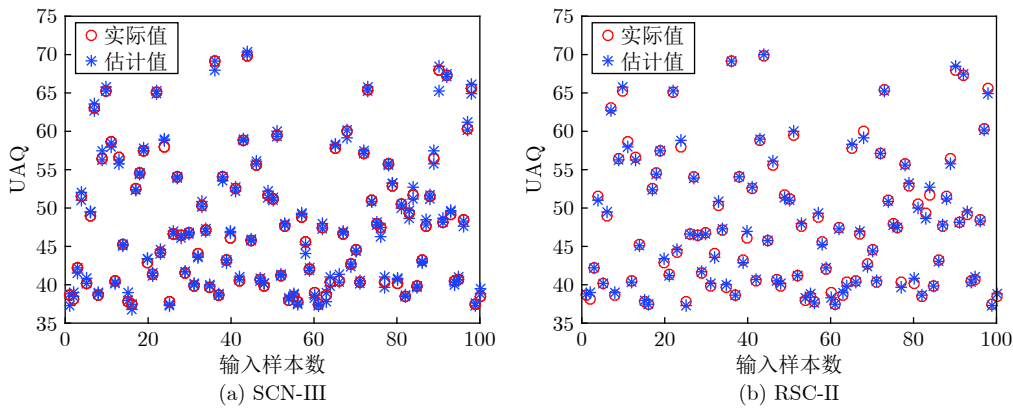


图 7  $L_{max} = 50$  所对应的 MFSP 数据集的测试性能: (a) SCN-III; (b) RSC-II  
Fig.7 Test performance at  $L_{max} = 50$  on the actual MFSP dataset: (a) SCN-III; (b) RSC-II

SCN 算法存在的过拟合问题。同时, RSC-II 算法具有与 SCN 算法相同的小尺寸特性。因此, 基于所提算法建立的模型具有尺寸小、鲁棒性强、泛化性能好的优点, 能够应用到实际主通风机切换过程进行井下供给风量估计。

### References

- 1 杨校卫, 韩传功, 何胜良, 等. 一种煤矿老通风系统不停风倒机装置, 中国专利CN208456655U, 2019年2月1日
- 2 Chatterjee A, Zhang L J, Xia X H. Optimization of mine ventilation fan speeds according to ventilation on demand and time of use tariff. *Applied Energy*, 2015, **146**: 65–73
- 3 Wu X Z, Ma X P, Ren Z H. Study on coal mine main fan automatic switchover aiming at ventilation unceasing and its numerical simulation. In: Proceedings of the 2nd International Asia Conference on Informatics in Control, Automation and Robotics. Wuhan, China: IEEE, 2010. 2: 227–230
- 4 Ge H Q, Ma X P, Wu X Z, Liang C. Study on mine main fan switchover aiming at invariant ventilation based on nonlinear constrained programming. *International Journal of Digital Content Technology and Its Applications*, 2012, **6**(17): 496–505
- 5 Wang Qian-Jin, Dai Wei, Yang Chun-Yu, Ma Xiao-Ping. Modeling and analysis of coal mine main fan switchover system. *Journal of China Coal Society*, 2018, **43**(S2): 606–614 (王前进, 代伟, 杨春雨, 马小平. 煤矿主通风机切换系统建模与分析. *煤炭学报*, 2018, **43**(S2): 606–614)
- 6 Wang Q J, Ma X P, Yang C Y, Dai W. Modeling and control of mine main fan switchover system. *ISA Transactions*, 2019, **85**: 189–199
- 7 Ren Zi-Hui, Wang Cui, Ni Ting-Ting, Cheng Jiang-Yang. Research on air volume of reversing without stopping ventilation based on neural network. *Coal Technology*, 2016, **35**(11): 229–231 (任子晖, 王翠, 倪婷婷, 成江洋. 基于神经网络不停风倒机风量变化的研究. *煤炭技术*, 2016, **35**(11): 229–231)
- 8 Zou W D, Xia Y Q, Li H F. Fault diagnosis of tennessee-castman process using orthogonal incremental extreme learning machine based on driving amount. *IEEE Transactions on Cybernetics*, 2018, **48**(12): 3403–3410
- 9 Han Min, Li De-Cai. An norm 1 regularization term ELM algorithm based on surrogate function and bayesian framework. *Acta Automatica Sinica*, 2011, **37**(11): 1344–1350 (韩敏, 李德才. 基于替代函数及贝叶斯框架的1范数ELM算法. *自动化学报*, 2011, **37**(11): 1344–1350)
- 10 Zou Wei-Dong, Xia Yuan-Qing. Virtual machine power prediction using incremental extreme learning machine based on compression driving amount. *Acta Automatica Sinica*, 2019, **45**(7): 1290–1297 (邹伟东, 夏元清. 基于压缩动量项的增量型ELM虚拟机能耗预测. *自动化学报*, 2019, **45**(7): 1290–1297)
- 11 Chen Xiao-Yun, Liao Meng-Zhen. Dimensionality reduction with extreme learning machine based on sparsity and neighborhood preserving. *Acta Automatica Sinica*, 2019, **45**(2): 325–333 (陈晓云, 廖梦真. 基于稀疏和近邻保持的极限学习机降维. *自动化学报*, 2019, **45**(2): 325–333)
- 12 Li M, Wang D H. Insights into randomized algorithms for neural networks: Practical issues and common pitfalls. *Information Sciences*, 2017, **382**: 170–178
- 13 Wang D H, Li M. Stochastic configuration networks: Fundamentals and algorithms. *IEEE Transactions on Cybernetics*, 2017, **47**(10): 3466–3479
- 14 Wang Qian-Jin, Ma Xiao-Ping, Zhang Shou-Tian. Application of PLC soft redundancy in monitoring and control system for ventilator. *Industry and Mine Automation*, 2014, **40**(1): 93–96 (王前进, 马小平, 张守田. PLC软冗余在通风机监控系统中的应用. *工矿自动化*, 2014, **40**(1): 93–96)
- 15 Ge H Q, Ma X P, Wu X Z, Zhang J. Fuzzy PID control of mine main fan switchover aiming at invariant ventilation. In: Proceedings of the 2011 International Conference on Intelligence Science and Information Engineering. Wuhan, China: IEEE, 2011. 325–328
- 16 Wang Q J, Dai W, Ma X P, Yang C Y. Multiple models and neural networks based adaptive PID decoupling control of mine main fan switchover system. *IET Control Theory & Applications*, 2018, **12**(4): 446–455
- 17 Ge H Q, Xu G, Huang J X, Ma X P. A mine main fans switchover system with lower air flow volatility based on improved particle swarm optimization algorithm. *Advances in Mechanical Engineering*, 2019, **11**(3): 1–10
- 18 Ray A L, Couse D. Cement plant fan efficiency upgrades. *IEEE Transactions on Industry Applications*, 2017, **53**(2): 1562–1568
- 19 Papar R, Szady A, Huffer W D, Martin V, McKane A. *Increasing Energy Efficiency of Mine Ventilation Systems*. Industrial Energy Analysis, Lawrence Berkeley National Laboratory, University of California, 1999.
- 20 Zhu S, Chang W T, Liu D. Almost sure exponential stability of stochastic fluid networks with nonlinear control. *International Journal of Control*, 2018, **91**(2): 400–410
- 21 Camilleri R, Howey D A, McCulloch M D. Predicting the temperature and flow distribution in a direct oil-cooled electrical machine with segmented stator. *IEEE Transactions on Industrial Electronics*, 2016, **63**(1): 82–91
- 22 Abareshi M, Hosseini S M, Sani A A. A simple iterative method for water distribution network analysis. *Applied Mathematical Modelling*, 2017, **52**: 274–287
- 23 Gorban A N, Tyukin I Y, Prokhorov D V, Sufeikov K I. Approximation with random bases: Pro et Contra. *Information Sciences*, 2016, **364**: 129–145
- 24 Lichman M. UCI machine learning repository [Online] available: <http://archive.ics.uci.edu/ml>, January 1, 2013



**王前进** 盐城工学院电气工程学院讲师。2018 年获得中国矿业大学博士学位。主要研究方向为数据驱动建模与控制, 机器学习算法。

E-mail: wangqianjinabc@163.com

(**WANG Qian-Jin** Lecturer at the School of Electrical Engineering,

Yancheng Institute of Technology. He received his Ph. D. degree from China University of Mining and Technology in 2018. His research interest covers data-driven modeling and control, and machine learning algorithm.)



杨春雨 中国矿业大学信息与控制工程学院教授. 2009 年获得东北大学博士学位. 主要研究方向为广义系统和鲁棒控制.

E-mail: chunyuyang@cumt.edu.cn

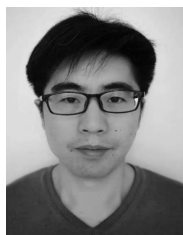
(**YANG Chun-Yu** Professor at the School of Information and Control Engineering, China University of Mining and Technology. He received his Ph. D. degree from Northeastern University in 2009. His research interest covers descriptor systems and robust control.)



马小平 中国矿业大学信息与控制工程学院教授. 主要研究方向为过程控制, 网络控制系统, 故障诊断. 本文通信作者.

E-mail: xpma@cumt.edu.cn

(**MA Xiao-Ping** Professor at the School of Information and Control Engineering, China University of Mining and Technology. His research interest covers process control, networked control system, and fault detection. Corresponding author of this paper.)



张春富 盐城工学院电气工程学院副教授. 2007 年获得哈尔滨工业大学博士学位. 主要研究方向为自动化测控.

E-mail: zhangchunfu@hit.edu.cn

(**ZHANG Chun-Fu** Associate professor at the School of Electrical Engineering, Yancheng Institute of Technology. He received his Ph. D. degree from Harbin Institute of Technology in 2007. His research interest covers automatic test and control.)



彭思敏 盐城工学院副教授, IEEE 高级会员. 2013 年获得上海交通大学电力系统及其自动化专业博士学位. 主要研究方向为新能源发电, 电池储能系统及微电网控制技术.

E-mail: psmsteven@163.com

(**PENG Si-Min** Associate professor at Yancheng Institute of Technology, IEEE senior member. He received his Ph. D. degree in electric power system and its automation from Shanghai Jiao Tong University in 2013. His research interest covers control of renewable sources generation, battery energy storage system and microgrid.)