

DOI: 10.11992/tis.201508015
网络出版地址: <http://www.cnki.net/kcms/detail/23.1538.TP.20160824.0928.002.html>

堆叠隐空间模糊 C 回归算法及其在 发酵数据多模型建模中的应用

刘欢, 王骏, 邓赵红, 王士同
(江南大学 数字媒体学院, 江苏 无锡 214122)

摘要:切换回归算法 FCR 的性能容易受到噪声点以及离群点的影响,同时该算法对于复杂数据的处理能力较差。对此,文中提出一种基于堆叠隐空间的模糊 C 回归算法。该算法将基于 ELM 特征映射技术,利用主成分分析进行特征提取,再结合多层前馈神经网络学习结构对隐空间进行多次扩展和压缩。实验结果表明,该算法具有更好的抗噪性能,对模糊指数的变化不敏感,同时在处理复杂数据以及在多模型建模中更加精确、高效、稳定。
关键词:隐空间映射;极限学习机;主成分分析;模糊 C 回归算法;多层神经网络;多模型建模
中图分类号:TP181 **文献标志码:**A **文章编号:**1673-4785(2016)05-0670-10

中文引用格式:刘欢,王骏,邓赵红,等.堆叠隐空间模糊 C 回归算法及其在发酵数据多模型建模中的应用[J]. 智能系统学报, 2016, 11(5): 670-679.
英文引用格式:LIU Huan, WANG Jun, DENG Zhaohong, et al. A cascaded hidden space fuzzy C-regression algorithm and its application in multi-model modeling for the fermentation process [J]. CAAI transactions on intelligent systems, 2016, 11(5): 670-679.

A cascaded hidden space fuzzy C-regression algorithm and its application in multi-model modeling for the fermentation process

LIU Huan, WANG Jun, DENG Zhaohong, WANG Shitong
(School of Digital Media, JiangNan University, WuXi 214122, China)

Abstract: The switching regression algorithm FCR is sensitive to noise data and outliers. The algorithm also has low levels of capability for dealing with complex data. In order to handle these problems, an improved fuzzy C-regression algorithm is proposed based on cascaded hidden space. In our method, principal component analysis is combined with extreme machine learning feature mapping and multilayer feedforward neural networks. The experimental results show that our proposed method is more stable as regards noise data and outliers, and thus more suitable for handling complex data and multi-model modeling problems for the fermentation process.
Keywords: hidden space feature mapping; extreme learning machine; principal component analysis; fuzzy C-regression algorithm; multilayer neural network; multi-model modeling.

切换回归模型在经济、数据挖掘等领域中有着广泛的应用。近年来,基于模糊聚类分析^[1-4]的切换回归技术得到了广泛的研究。如文献[5]在模糊 C 均值聚类算法(fuzzy C clustering algorithm, FCM)的基础上提出了模糊 C 均值回归算法(fuzzy C re-

gression algorithm, FCR)算法,从而实现了对多个模型的切换回归;在此基础上,文献[6]结合牛顿万有引力定律提出了关于切换回归的集成模糊聚类算法(gravity-based fuzzy clustering algorithm, GFC)算法,文中实验表明该算法与 FCR 相比具有更快的收敛速度;文献[7]通过向 FCR 中引入动态的样本权值,提出了离群模糊切换回归模型,从而使模型具有更强的鲁棒性,并且能通过权值发现数据集中的离群点。虽然这些方法在一定条件下可以较好地解决切

收稿日期:2015-08-14. 网络出版日期:2016-08-24.
基金项目:国家自然科学基金项目(61300151);江苏省自然科学基金项目(BK20130155, BK20130160).
通信作者:刘欢.E-mail: 771627297@qq.com.

换回归模型中的参数学习问题,但是面向现实生活中的复杂数据,仍然有诸多局限性,其学习能力有待进一步提高。

近年来,以极速学习机(extreme learning machine, ELM)为代表的单隐层前馈神经网络快速学习理论得到了研究人员的深入研究^[8-11]。研究表明,将 ELM 特征映射技术代替已有核方法中的核映射,可以有效提高学习器的学习能力,目前该技术已被广泛用于分类、回归、聚类等学习任务中。结合已有的研究工作,本文重点研究基于 ELM 隐空间学习理论的切换回归模型。首先研究基于主成分分析(principal component analysis, PCA)^[12]的压缩隐空间构建新方法。在此基础上,结合多层神经网络学习方法,将单隐层结构改造为多隐层结构^[13],提出堆叠隐空间模糊 C 回归算法(cascaded hidden space FCR, CHS-FCR)。该方法通过使用层次化的学习结构对数据对象在不同层次上的表达形式进行抽象,并通过重组低层概念来重新定义高层概念,从而有效提高了学习系统处理复杂问题的能力。经实验验证,该算法能有效地弥补经典 FCR 的若干不足,在保证学习精度和学习效率的前提下,该方法对噪声数据和离群点有很好的鲁棒性。

1 相关工作

1.1 ELM 隐空间

在 ELM 中,隐节点所形成的特征空间构成隐空间^[14]。其映射过程如下:1)随机生成权重矩阵 $\mathbf{W} \in \mathbf{R}^{L \times d}$ 和偏移量矩阵 $\mathbf{B}=[b_1 \ b_2 \cdots b_L]^T$,其中 L 是 ELM 隐节点总数, d 是原始数据的维数。2)将原始数据映射到 L 维的隐空间中。每一个输入数据都是一个 d 维的向量, $\mathbf{x}=[x_1 \ x_2 \cdots x_d]^T$ 。该特征映射可以表示为:

$$\mathbf{h}(\mathbf{x}) = [h_1(\mathbf{x}) \ \cdots \ h_i(\mathbf{x}) \ \cdots \ h_L(\mathbf{x})]^T =$$
$$[G(w_1, b_1, \mathbf{x}) \ \cdots \ G(w_i, b_i, \mathbf{x}) \ \cdots \ G(w_L, b_L, \mathbf{x})]^T \quad (1)$$

式中 $G(\mathbf{x})$ 是激励函数,其映射过程如图 1 所示。

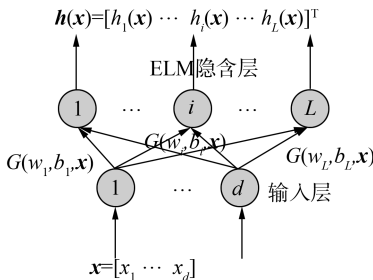


图 1 隐空间特征映射的过程

Fig.1 The process of hidden-space feature mapping

常用的激励函数有以下几种:

Sigmoid:

$$G(x) = \frac{1}{1 + \exp(-x)} \quad (2)$$

Sine:

$$G(x) = \sin(x) \quad (3)$$

Gaussian:

$$G(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right) \quad (4)$$

1.2 模糊 C 回归算法及其问题分析

1993 年 Richard J. Hathaway^[5] 和 James C. Bezdek^[15] 提出了 FCR 算法。该算法能在对观察数据进行模糊划分的同时估计出划分数据满足的回归模型参数。设数据集 $D = \{(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)\}$, 其中 n 是数据点个数, $x_j = (x_{j1}, x_{j2}, \cdots, x_{jd}) \in \mathbf{R}^d, y_j \in \mathbf{R}$ 表示观察数据模型, d 是数据集的特征数。FCR 算法构建的回归模型为

$$y_{j,i} = \beta_{i0} + \beta_{i1}x_{j1} + \cdots + \beta_{id}x_{jd}$$
$$i = 1, 2, \cdots, c; j = 1, 2, \cdots, n \quad (5)$$

式中 c 是模型数目, j 是样本点个数, β_i 为模型参数。FCR 算法的目标函数为

$$J_{\text{FCR}} = \sum_{i=1}^c \sum_{j=1}^n u_{ij}^m (y_j - y_{j,i})^2 \quad (6)$$

式中 u_{ij} 是隶属度,且满足 $\sum_{i=1}^c u_{ij} = 1, u_{ij} \in \{0, 1\}, i = 1, 2, \cdots, c; j = 1, 2, \cdots, n$ 。 m 是模糊指数且 $m > 1$ 。 FCR 使用式(7)来更新隶属度:

$$u_{ij} = \frac{(y_j - y_{j,i})^{-2/(m-1)}}{\sum_{k=1}^c (y_j - y_{j,k})^{-2/(m-1)}} \quad k = 1, 2, \cdots, c \quad (7)$$

FCR 的目标函数是一个基于模糊划分的多模型最小二乘拟合准则问题,任何现有的能解决加权最小二乘问题的方法都可以用来估计参数。模型参数可以通过式(8)求解:

$$\boldsymbol{\beta}_i = \begin{bmatrix} \beta_{i0} \\ \vdots \\ \beta_{id} \end{bmatrix} = [\mathbf{X}^T \mathbf{D}_i \mathbf{X}]^{-1} \mathbf{X}^T \mathbf{D}_i \mathbf{Y} \quad (8)$$

式中 $\mathbf{X} \in \boldsymbol{\Omega} \subset \mathbf{R}^{n \times (d+1)}$ 是以 $(1, x_j) = (1, x_{j1}, \cdots, x_{jd})$ 为行向量输入数据矩阵。 $\mathbf{Y} \in \boldsymbol{\Omega} \subset \mathbf{R}^n$ 是以输出 y_j 为行向量的输出列向量; $\mathbf{D}_i \in \boldsymbol{\Omega} \subset \mathbf{R}^{n \times n}$ 是以 u_{ij}^m 第 j 个对角元素的对角矩阵, $\mathbf{D}_i = \text{diag}(u_{i1}^m, u_{i1}^m, \cdots, u_{in}^m)$ 。 FCR 算法的步骤为:

- 1) 给定模型参数 $c(1 \leq c \leq n)$,模糊指数 $m > 1$,迭代终止条件 $\varepsilon > 0$,并初始化划分矩阵 $\mathbf{U}^{(0)}$,迭代步骤 $l = 0$;
- 2) 根据式(8)计算模型参数 $\boldsymbol{\beta}_i^l$;
- 3) 根据 $(y_j - y_{j,i})^2$ 计算拟合误差,代入式(7)求隶属度矩阵 \mathbf{U} ;
- 4) 根据式(6)计算目标函数值,若

$\|J_{\text{FCR}}^{(l+1)} - J_{\text{FCR}}^{(l)}\| \leq \varepsilon$ 则迭代终止, 其中 ε 表示迭代终止的阈值; 否则 $l=l+1$, 转至步骤 2)。

研究表明, FCR 可以较好地解决切换回归问题, 但是也存在以下问题: 对噪声和离群点敏感, 当数据集中含有离群点或大量噪声数据时, 算法性能就会受到影响; FCR 算法在数据原空间中进行学习, 针对现实生活中的复杂数据, FCR 算法的性能很难得到有效地提升。

本文将 ELM 隐空间学习理论与特征降维技术相结合, 基于多隐层神经网络学习方法, 将传统的基于浅层结构的学习方法改造为多隐层学习结构, 从而提高对复杂数据的无监督学习能力。通过提出堆叠隐空间学习结构, 从而有效精简冗余信息和过滤噪声数据, 同时补充必要的信息, 使得切换回归算法具有更强的鲁棒性, 并将其应用于多模型建模以取得更好的效果。

2 堆叠隐空间模糊 C 回归算法

2.1 基于主成分分析的压缩 ELM 隐空间

在 ELM 学习技术中, 可以通过随机赋值的方法来快速生成 ELM 隐空间, 随着隐节点数目的增加, 学习精度也不断提高^[16-17]。但是随之而来的一个重要问题是计算效率会逐步降低。此外, 由于隐节点生成过程中相关参数是随机生成的, 因此不可避免地会引入大量噪声。针对这一问题, 本文中 will 将 ELM 隐空间技术与主成分分析相结合, 提出基于主成分分析的压缩隐空间构建方法, 其过程如下:

1) 根据图 1 所示过程将原始数据映射到高维空间 R^L 中, 得到矩阵 $H^{(1)} \in R^{n \times L}$, n 为样本点个数;

2) 对 $H^{(1)}$ 的每一列 $H_j^{(1)}, j=1, 2, \dots, L$ 进行归一化, 得到矩阵 $A=[A_1 A_2 \dots A_L]$, 其中 A_j 计算方法为

$$A_j = H_j^{(1)} - \frac{1}{n} \sum_{j=1}^N H_j^{(1)}, j=1, 2, \dots, L \quad (9)$$

3) 计算矩阵 A 的协方差矩阵 C , 并计算矩阵 C 的特征值 λ_i 和特征向量 $V_i, i=1, 2, \dots, L$;

4) 将特征值 λ_i 降序排序, 取前 p 个特征值对应的特征向量, 按列组成矩阵 \tilde{V} , 特征提取后的矩阵为

$$H^{(2)} = A\tilde{V} \quad (10)$$

其结构可映射为如图 2 所示的前馈神经网络模型。输入数据经过输入层后, 通过 ELM 特征映射被映射到高维 ELM 隐空间中。在此过程中, 生成 ELM 隐空间时的随机赋值操作会引入噪声, 后续的 PCA 对高维隐空间中的数据进行降维, 从而有效地过滤掉部分噪声, 这有利于提高学习性能。

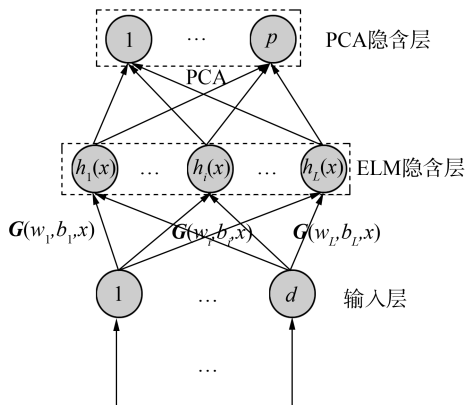


图 2 基于 PCA 的压缩 ELM 隐空间

Fig.2 Cascade hidden space based on PCA

2.2 堆叠隐空间模糊 C 回归算法 CHS-FCR

在隐空间构建过程中, 为了使学习器得到更好的表达能力和稳定的学习效果, 通常会使用较多的隐节点数目。但是, 这增加了额外的计算负担。研究表明, 在保证学习器泛化能力的前提下, 将单隐层结构改造为多隐层结构是降低隐节点数目的有效方法。为此, 本文对上述基于 PCA 的压缩隐空间映射过程进行改造, 通过把 ELM 隐层中的隐节点分散到多个隐含层中, 并与 PCA 层的隐结点相结合, 形成新的混合隐含层。将若干个混合隐含层进行叠加, 得到多隐层学习结构, 该过程如图 3 所示。

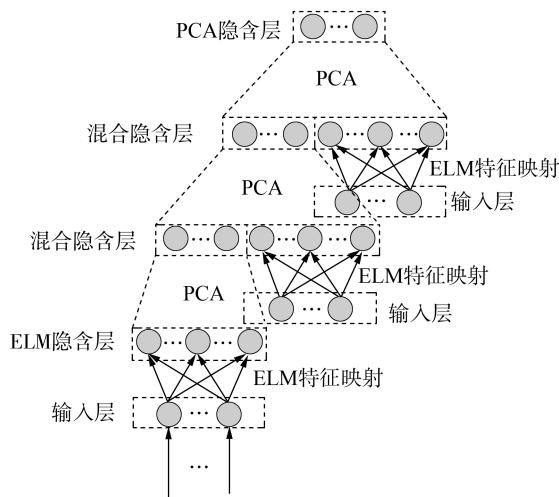


图 3 堆叠隐空间学习结构

Fig.3 Cascade hidden space learning structure

堆叠隐空间模糊 C 回归算法 CHS-FCR 的描述如下:

设隐节点总数为 L , 数据集 D 的维数为 d , PCA 压缩后的维数为 p , 隐空间压缩的次数为 f , 每层隐空间的 ELM 隐节点数 $T=L/f, n$ 为样本点个数。

1) ① 随机生成值在 $[0, 1]$ 之间的权重矩阵

$W \in \mathbf{R}^{T \times d}$ 以及值在 $[0.5, 1]$ 之间的偏移量矩阵 $B \in \mathbf{R}^{T \times 1}$ 。

②根据式(1)以及图 1 将数据集 D 进行特征映射,得到高维数据矩阵 $H^{(1)} \in \mathbf{R}^{n \times T}$ 。

2) 利用基于 PCA 的压缩隐空间方法对 $H^{(1)}$ 进行压缩,得到维数为 p 的数据矩阵 $H^{(2)}$;

3) for $j = 1:f-1$

①重复步骤 1 中 1)、2);

②将 $H^{(2)}$ 与 $H^{(1)}$ 合并存入矩阵 $H^{(3)}$ 中, $H^{(3)} = [H^{(2)} H^{(1)}]$ 。

③利用 PCA 对 $H^{(3)}$ 进行特征提取,得到矩阵记为 $H^{(2)}$;

循环结束;

4) 构造全一矩阵 $\mathbf{1} \in \mathbf{R}^{n \times 1}$, 将最终获取的数据 $H^{(2)} = [\mathbf{1}, H^{(2)}]$ 利用 FCR 框架进行切换回归。

CHS-FCR 算法将基于 PCA 的压缩 ELM 隐空间单隐层学习结构改造为多隐层学习结构。假设从第二层开始,每个隐含层中随机生成的 ELM 隐节点数都是 T ,由前一层 PCA 压缩后得到的隐节点数为 T' ,易得隐节点总数为

$$L = (T + T') \times (f - 1) + T \tag{11}$$

其中随机生成的 ELM 隐节点总数为

$$L_{\text{ELM}} = T \times f \tag{12}$$

经过多次隐空间压缩,噪声信息被过滤,同时会发生相应的信息损失;但是每个混合层中新生成的 ELM 隐节点信息弥补了这些丢失的信息,因此本文的堆叠隐空间结构能得到更好的学习效果。

3 实验研究与分析

3.1 实验平台以及算法性能的评价指标

本文在表 1 中实验平台上进行所有实验并且采用式(13)和(14)所示的指标来评价各算法的性能。

$$J_{\text{RRSE}} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i' - y_i)^2 / \frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^2} \tag{13}$$

$$J_{\text{SCC}} = \frac{(N \sum_{i=1}^N y_i' \cdot y_i)^2}{[N \sum_{i=1}^N (y_i')^2 - (\sum_{i=1}^N y_i')^2] \cdot [N \sum_{i=1}^N (y_i)^2 - (\sum_{i=1}^N y_i)^2]} \tag{14}$$

式中: N 表示测试样本数, y_i 表示测试样本第 i 次输入的输出, y_i' 表示测试样本第 i 次输入的模糊系统输出, $\bar{y} = \frac{1}{N} \sum_{i=1}^N y_i$ 。式(13)和(14)所示的指标数值分别越接近于 0 和 1,则算法的性能越好。

表 1 实验平台

Table 1 Experimental platform	
结构	配置
CPU	Intel(R) Core(TM) i5-4590 3.30GHz
内存	8.00 GB
操作系统	64bit Windows7
软件平台	Matlab R2012a

3.2 鲁棒性实验

本实验中,基于文献[5]的模拟回归数据集,通过分别加入噪声点和离群点来比较 FCR 以及 CHS-FCR 算法的鲁棒性。实验中,FCR 算法和 CHS-FCR 算法的模型个数 $c=2$,模糊指数 $m=2$,CHS-FCR 隐空间压缩次数 $f=5$,PCA 特征提取后的维数 $p=3$ 。

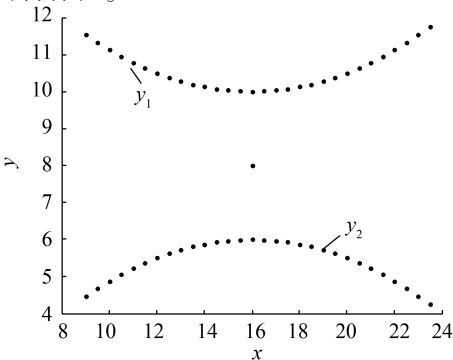
3.2.1 离群点实验

在切换回归分析中,离群点主要是指不符合任一模型的数据,它主要是在收集数据的过程中出现误操作或异常情况而引入的。本实验中基于式(15)和(16)生成 2 个回归数据集,并分别加入离群点(16,8)。实验数据集分布如图 4 所示。

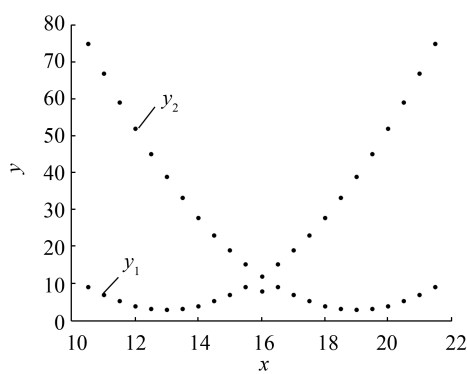
$$A: \begin{cases} y_1 = 18 - x + 0.031\ 25x^2 \\ y_2 = -2 + x - 0.031\ 25x^2 \end{cases} \tag{15}$$

$$B: \begin{cases} y_1 = 172 - 26x + x^2 \\ y_2 = 364 - 38x + x^2 \end{cases} \tag{16}$$

图 5 和图 6 给出了加入离群点之前模糊 C 均值回归算法(fuzzy C regression algorithm, FCR)和堆叠隐空间模糊 C 回归算法(cascaded hidden space FCR, CHS-FCR)的回归结果,图 7 和图 8 给出了加入离群点之后两算法的回归结果。从中可以看出,加入离群点之前,FCR 和 CHS-FCR 这 2 个算法均可以得到理想的结果。加入离群点之后,FCR 的拟合效果明显受到离群点的影响;但是 CHS-FCR 仍然能得到令人满意的拟合结果。可见,压缩隐空间方法的引入使本文所提的 CHS-FCR 算法对离群点具有更好的鲁棒性。



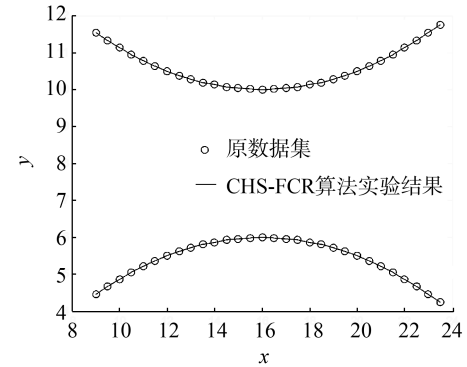
(a)数据集 A



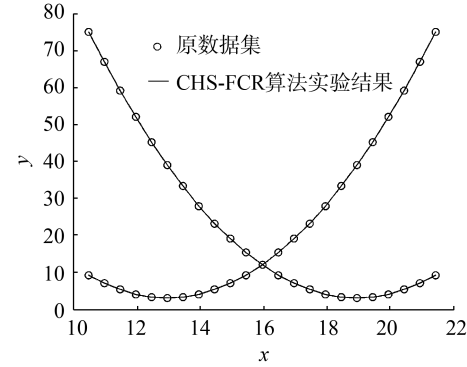
(b) 数据集 B

图 4 加入离群点后的数据及分布

Fig.4 Data distribution after adding outliers



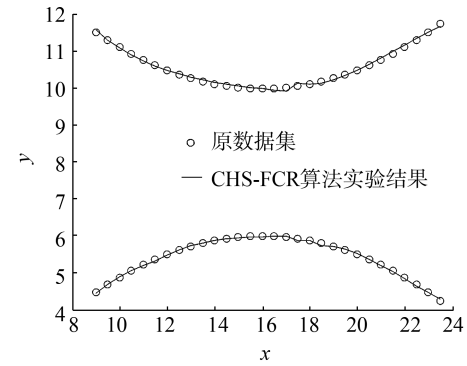
(a) 数据集 A



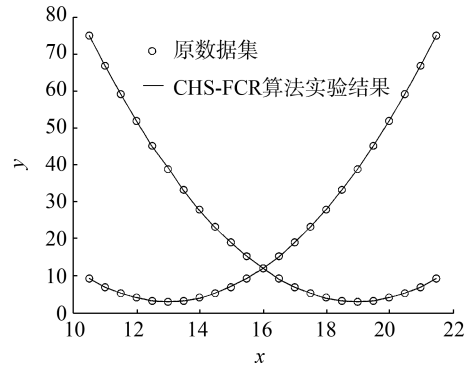
(b) 数据集 B

图 5 FCR 算法原数据集测试结果

Fig.5 FCR algorithm test results of original dataset



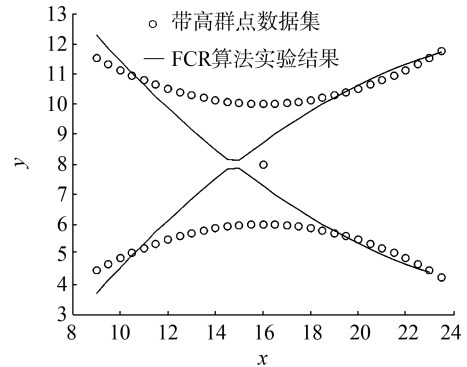
(a) 数据集 A



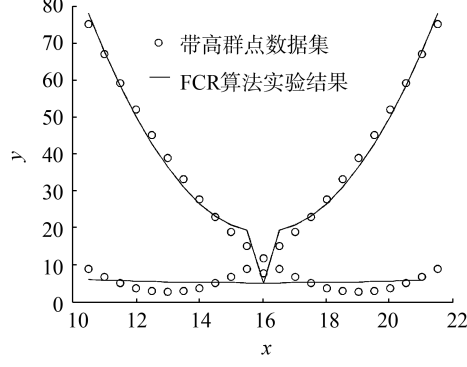
(b) 数据集 B

图 6 CHS-FCR 算法原数据集测试结果

Fig.6 CHS-FCR algorithm test results of original dataset



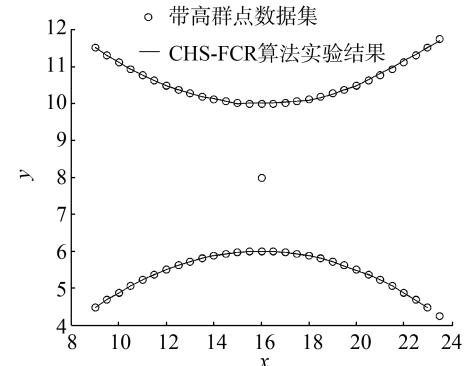
(a) 数据集 A



(b) 数据集 B

图 7 FCR 算法在离群点数据集上的测试结果

Fig.7 FCR algorithm test results of outliers dataset



(a) 数据集 A

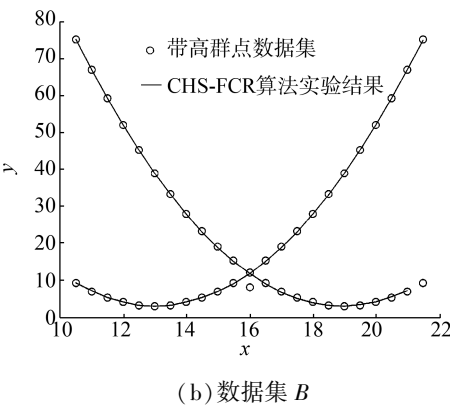


图 8 CHS-FCR 算法在离群点数据集上的测试结果

Fig.8 CHS-FCR algorithm test results of outliers dataset

3.2.2 抗噪声实验

为了进一步验证本文算法的抗噪性能并将其与 FCR 进行比较,本实验采用式(17)生成带随机噪声的数据集。

$$C: \begin{cases} y_1 = 18 - x + 0.031\,25x^2 + \delta \\ y_2 = -2 + x - 0.031\,25x^2 + \delta \end{cases} \quad (17)$$

式中 $\delta \in [-0.5, 0.5]$ 为服从均匀分布的随机值。分别运行 FCR 和 CHS-FCR, 所得结果如图 9 所示。从图 9 的模拟实验结果中可以看出, 在模拟数据集 C 中, FCR 算法易受噪声点的影响, 无法得到准确的实验结果。相反, CHS-FCR 算法在该数据集中能够取得较好的拟合效果。

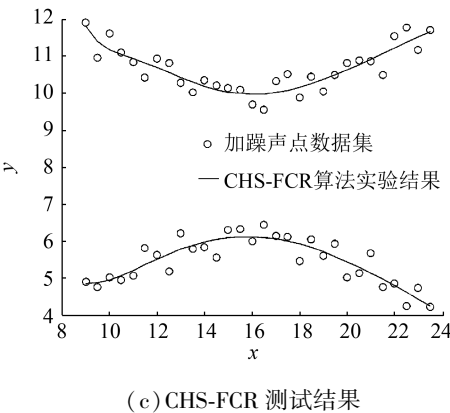
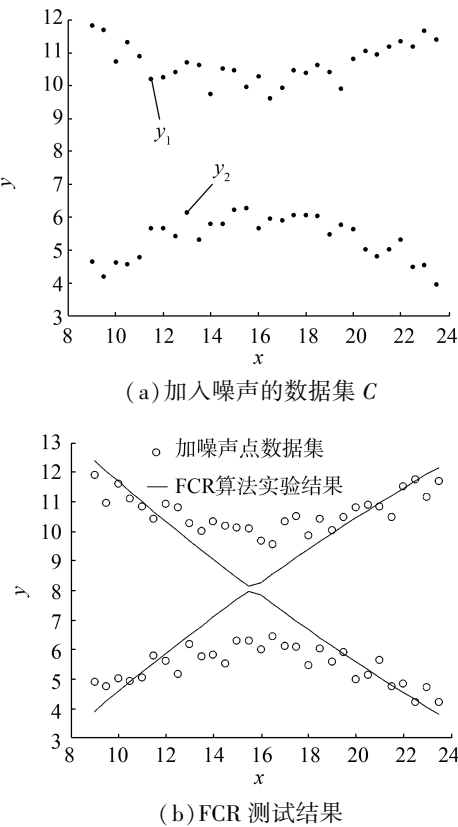


图 9 FCR 以及 CHS-FCR 算法噪声数据集 C 上的测试结果

Fig.9 FCR and CHS-FCR algorithms test results of noise dataset

综合以上 2 个实验不难发现, 传统的 FCR 算法性能易受离群点和噪声数据的干扰, 而 CHS-FCR 算法利用多次隐空间映射和压缩, 使得隐空间中的冗余信息被精简, 噪声被有效过滤, 同时每一层中进行了适度的信息弥补, 这使得该算法在不同回归模型中均能很好地完成学习过程。

3.3 发酵数据集多模型建模

切换回归模型由多个简单回归模型混合而成, 本文所研究的切换回归模型可用于发酵数据集多模型建模。为了更好地表明多层压缩隐空间对回归性能的影响, 在保持 ELM 隐节点数目不变的前提下, 给出了 CHS-FCR 的 2 个不同版本: CHS-FCR ($f=1$) 和 CHS-FCR ($f=5$), 其中 f 为隐空间压缩的次数。通过使用 CHS-FCR ($f=5$)、FCR、HS-FCR 以及 CHS-FCR ($f=1$) 算法进行训练得到训练数据集中各样本点的隶属度 u_{ik} , $k=1, 2, \dots, n$; $i=1, 2, \dots, c$ 并使用式(18)计算各类聚类中心 v_i 。对于 CHS-FCR($f=5$) 和 CHS-FCR($f=1$), 还分别记录各层中的 ELM 隐空间映射矩阵 W 、偏移量矩阵 B 和 PCA 过程中的变换矩阵 \tilde{V} 。在测试过程中, 通过式(1)、(9)、(10)将测试样本 x_i 映射到压缩隐空间中, 再由式(5)求出测试回归模型 $y_{t,i}$ 并使用式(18)、(19)得到各测试点的隶属度。预测值由 $y_t' = \sum_{i=1}^c u_{it}^m \cdot y_{t,i}$ 计算而得。

$$v_i = \frac{\sum_{k=1}^n u_{ik}^m h(x_i)}{\sum_{k=1}^n u_{ik}^m}, i = 1, 2, \dots, c \quad (18)$$

$$u_{it} = \frac{d_{it}^{-1/(m-1)}}{\sum_{j=1}^c d_{jt}^{-1/(m-1)}} \quad (19)$$

式中: $j=1,2,\cdots,c;d_{it}=\|\mathbf{v}_i-\mathbf{h}(x_t)\|^2$ 。

本文使用参考文献[18]中所用的发酵数据集并且采用10折交叉验证的方法进行实验。将CHS-FCR($f=5$)算法和FCR、HS-FCR、CHS-FCR($f=1$)算法进行比较。实验中,对于不同的数据集,采用不同的模糊指数 m 和模型个数 c ,并且让各算法都在同样的参数条件下运行。表2给出各数据集的详细信息及该数据集在实验中的相关参数设置,表3给出各算法说明及相关参数设置。

表 2 实验数据集

Table 2 Experimental datasets					
数据集 序号	数据集	样本数	特征 数	模糊 指数 m	模型 个数 c
D_1	bio_P_scal_P ₁	1 002	6	2	5
D_2	bio_S_scal_P ₁	1 002	6	3	5
D_3	bio_P_scal_P ₃	3 004	6	4	5
D_4	bio_S_scal_P ₃	3 004	6	4	25

3.3.1 算法的拟合性能对比实验

本实验从拟合精度出发来研究CHS-FCR($f=5$)、CHS-FCR($f=1$)与FCR以及HS-FCR在表2所示中的发酵数据集上的多模型建模效果。实验中,

表 4 各算法在不同数据集上的拟合精度对比

Table 4 Each algorithm compared fitting accuracy on different datasets					
数据集	性能指标	FCR	HS-FCR	CHS-FCR($f=1$)	CHS-FCR($f=5$)
D_1	J_{RRSE}	0.236 6±0.064 5	0.121 7±0.029 2	0.204 7±0.030 6	0.194 0±0.019 7
	J_{SCC}	0.963 9±0.016 7	0.985 3±0.006 8	0.962 1±0.014 1	0.969 6±0.010 4
D_2	J_{RRSE}	0.386 1±0.058 3	0.193 5±0.028 3	0.158 1±0.018 9	0.146 7±0.008 6
	J_{SCC}	0.925 3±0.010 2	0.964 5±0.005 4	0.975 8±0.005 8	0.979 9±0.003 7
D_3	J_{RRSE}	0.331 5±0.035 2	0.208 7±0.043 5	0.205 2±0.037 8	0.204 2±0.030 1
	J_{SCC}	0.930 9±0.011 7	0.959 3±0.020 7	0.958 8±0.008 7	0.969 1±0.006 3
D_4	J_{RRSE}	0.711 8±0.076 7	0.145 8±0.021 3	0.154 5±0.021 4	0.144 7±0.017 4
	J_{SCC}	0.843 1±0.028 4	0.980 0±0.006 3	0.978 0±0.005 0	0.980 8±0.004 0

3.3.2 算法效率对比实验

本实验研究在相同ELM隐节点总数的前提下,浅层学习结构和多层学习结构对算法效率的影响。实验中分别执行CHS-FCR($f=5$)与HS-FCR算法,

分别运行各算法10次,得到 J_{RRSE} 、 J_{SCC} 这2个指标的均值和标准差如表4所示。

表 3 各算法的说明以及相关参数设置

Table 3 The description of each algorithms and associated parameters	
算法	算法说明以及相关参数
FCR	最大迭代次数100次,收敛阈值 10^{-5} 。
HS-FCR	通过ELM特征映射后进行执行FCR算法,ELM隐节点总数1 000,激励函数Sigmoid,最大迭代次数100次,收敛阈值 10^{-5} 。
CHS-FCR($f=1$)	通过ELM特征映射后利用PCA进行1次隐空间压缩后执行FCR算法,ELM隐节点总数1 000,PCA提取的特征数为5,最大迭代次数100次,收敛阈值 10^{-5} 。
CHS-FCR($f=5$)	通过ELM特征映射后利用PCA进行5次隐空间压缩后执行FCR算法,ELM隐节点总数1 000,PCA提取的特征数为5,最大迭代次数100次,收敛阈值 10^{-5} 。

从表4的实验结果可以发现,在发酵数据集中CHS-FCR($f=5$)与FCR和HS-FCR这2个算法相比拥有更好的学习效果。将CHS-FCR($f=5$)与CHS-FCR($f=1$)进行比较,不难发现,将ELM隐节点分散到多个混合隐含层中,经过多次隐空间压缩,有助于进一步提高算法的学习精度和算法的稳定性。

并为2个算法设置相同的ELM隐节点总数目为500、1 000。分别使2个算法各运行10次,记录 J_{RRSE} 、 J_{SCC} 以及算法运行时间这3个指标的均值如表5所示。由表5中的HS-FCR的运行结果可以看出,

通过 ELM 映射将原数据经映射到高维 ELM 隐空间使得隐节点总数增加,后续学习过程的效率明显降低。而本文所提出的 CHS-FCR ($f=5$) 中,虽然涉及了多次隐空间映射和压缩,但是在隐空间压缩过程中,冗余信息被有效精简,这使得 CHS-FCR($f=5$) 算

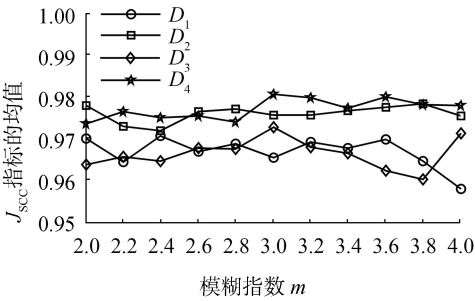
法能够在高效处理复杂数据的同时具有更好的拟合效果。由此可见,相同 ELM 隐节点总数的情况下,本文提出的具有多层学习结构堆叠隐空间构造技术可以使后续的回归算法取得更好的学习效果。

表 5 CHS-FCR($f=5$) 与 HS-FCR 算法的效率对比
Table 5 CHS-FCR($f=5$) compared with the efficiency of the HS-FCR

数据集	算法平均时间/s		ELM 隐节点总数		J_{RRSE} 平均值		J_{SCC} 平均值	
	HS-FCR	CHS-FCR($f=5$)	HS-FCR	CHS-FCR($f=5$)	HS-FCR	CHS-FCR($f=5$)	HS-FCR	CHS-FCR($f=5$)
D_1	62.136 8	1.370 0	500	500	0.122 7	0.231 8	0.984 2	0.960 6
	332.549 6	1.447 8	1 000	1 000	0.121 7	0.194 0	0.985 3	0.969 6
D_2	74.988 4	1.369 4	500	500	0.191 1	0.163 8	0.963 7	0.966 7
	352.609 2	1.702 4	1 000	1 000	0.193 5	0.146 7	0.964 5	0.979 9
D_3	148.269 0	15.731 2	500	500	0.211 2	0.209 7	0.959 7	0.960 1
	639.014 7	17.168 0	1 000	1 000	0.208 7	0.204 2	0.958 8	0.969 1
D_4	227.308 6	25.834 0	500	500	0.152 3	0.157 9	0.972 9	0.976 7
	750.463 8	27.408 3	1 000	1 000	0.145 8	0.144 7	0.980 0	0.980 8

3.3.3 参数敏感性实验

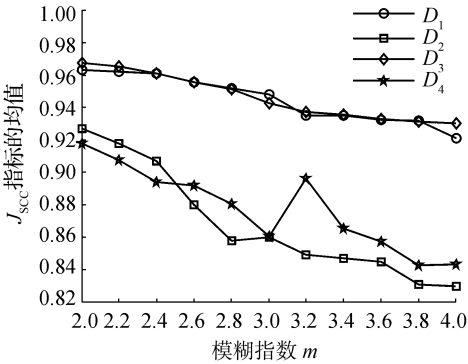
在 CHS-FCR($f=5$) 和 FCR 算法中,模糊指数 m 是一个重要指标,它严重影响着算法的执行效果。基于表 2 中的各数据集,本实验将研究模糊指数 m 的变化对算法性能产生的影响。给出了 J_{RRSE} 、 J_{SCC} 指标的变化情况,实验中, m 值在 $\{2, 2.2, 2.4, 2.6, 2.8, 3.0, 3.2, 3.4, 3.6, 3.8, 4.0\}$ 上变化,算法的其他参数如表 3 所示,2 个算法分别运行 10 次。记录 J_{RRSE} 和 J_{SCC} 指标的均值,实验结果如图 10、11 所示。



(b) CHS-FCR($f=5$)

图 10 FCR 和 CHS-FCR($f=5$) 算法的 J_{SCC} 指标随 m 的变化情况

Fig.10 J_{SCC} index of FCR and CHS-FCR($f=5$) algorithms with the change of m



(a) FCR

从图 10、11 结果中可以看出 FCR 算法的性能在 4 个数据集中随着 m 的变化有着较大的波动。例如在图 11(a) 中,当 $m=3.2$ 时 FCR 算法的 J_{RRSE} 指标会出现明显的波动现象,而图 11(b) 中的 CHS-FCR($f=5$) 算法无论其 m 值怎样变化,其 J_{RRSE} 指标基本呈现平稳变化的趋势。由此可见,虽然模糊指数 m 在较为宽广的范围内变化,但是 CHS-FCR

($f=5$)在各数据集上均可以取得令人满意且较稳定的结果。

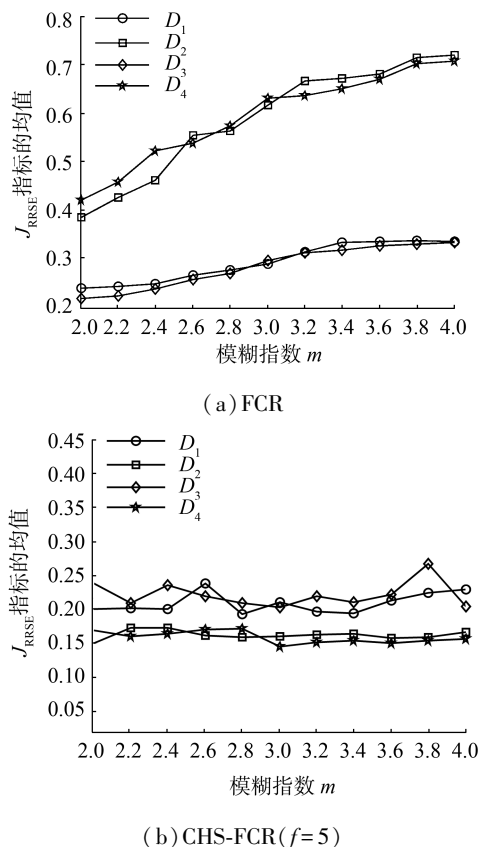


图 11 FCR 和 CHS-FCR ($f=5$) 算法的 J_{RRSE} 指标随 m 的变化情况

Fig.11 J_{RRSE} index of FCR and CHS-FCR ($f=5$) algorithms with the change of m

综上所述,经过堆叠隐空间的压缩与重组,CHS-FCR($f=5$)算法对模糊指数 m 的变化具有更好的鲁棒性,这有利于用户在实际应用中更方便地选取模糊指数。

4 结束语

本文基于主成分分析和 ELM 映射技术将复杂数据映射到低维空间中,并结合多层神经网络学习方法将单隐层结构改造为多隐层结构;在此基础上提出堆叠隐空间模糊 C 回归算法 CHS-FCR。通过在模拟以及发酵数据集的实验结果,均表明本文方法较其他相关算法有更好的鲁棒性且能够高效地处理复杂数据,可以有效地应用于发酵数据集多模型建模工作中。当前,随着回归模型算法研究的不断深入,人们已经基于不同理论提出了更先进的回归技术。如何将本文所提的堆叠隐空间技术应用于这些回归算法中,这将是今后研究的重点。

参考文献:

- [1] 王骏, 王士同, 邓赵红. 聚类分析研究中的若干问题[J]. 控制与决策, 2012, 27(3): 321-328.
WANG Jun, WANG Shitong, DENG Zhaohong. Survey on challenges in clustering analysis research[J]. Control and decision, 2012, 27(3): 321-328.
- [2] 蒋亦樟, 邓赵红, 王骏, 等. 熵加权多视角协同划分模糊聚类算法[J]. 软件学报, 2014, 25(10): 2293-2311.
JIANG Yizhang, DENG Zhaohong, WANG Jun, et al. Collaborative partition multi-view fuzzy clustering algorithm using entropy weighting[J]. Journal of software, 2014, 25(10): 2293-2311.
- [3] 张敏, 于剑. 基于划分的模糊聚类算法[J]. 软件学报, 2004, 15(6): 858-868.
ZHANG Min, YU Jian. Fuzzy partitional clustering algorithms[J]. Journal of software, 2004, 15(6): 858-868.
- [4] 秦蓓蓓. 基于聚类分析的鲁棒自适应切换回归算法研究[D]. 上海: 上海交通大学, 2012: 14-39.
QIN Beibei. Research on the robust and adaptive switching C-regressions models based on cluster analysis[D]. Shanghai: Shanghai Jiao Tong University, 2012: 14-39.
- [5] HATHAWAY R J, BEZDEK J C. Switching regression models and fuzzy clustering[J]. IEEE transactions on fuzzy systems, 1993, 1(3): 195-204.
- [6] WANG Shitong, JIANG Haifeng, LU Hongjun. A new integrated clustering algorithm GFC and switching regressions[J]. International journal of pattern recognition and artificial intelligence, 2002, 16(4): 433-446.
- [7] 沈红斌, 王士同, 吴小俊. 离群模糊切换回归模型研究[J]. 华东船舶工业学院学报: 自然科学版, 2003, 17(3): 31-36.
SHEN Hongbin, WANG Shitong, WU Xiaojun. Research on fuzzy switching regression models with outliers[J]. Journal of east China shipbuilding institute: natural science edition, 2003, 17(3): 31-36.
- [8] WANG Shitong, CHUNG Fulai, WANG Jun, et al. A fast learning method for feedforward neural networks[J]. Neurocomputing, 2015, 149: 295-307.
- [9] HUANG Guangbin, ZHU Qinyu, SIEW C K. Extreme learning machine: theory and applications[J]. Neurocomputing, 2006, 70(1/2/3): 489-501.
- [10] HUANG Guangbin, ZHOU Hongming, DING Xiaojian, et al. Extreme learning machine for regression and multiclass classification[J]. IEEE transactions on systems, man, and cybernetics, part b (cybernetics), 2012, 42(2): 513-529.
- [11] HUANG Guangbin, WANG Dianhui, LAN Yuan. Extreme

learning machines: a survey [J]. International journal of machine learning and cybernetics, 2011, 2(2): 107-122.

[12] JOLLIFFE I T. Principal component analysis [J]. New York: Springer, 2002.

[13] ZHOU Hongming, HUANG Guangbin, LIN Zhiping, et al. Stacked extreme learning machines [J]. IEEE transactions on cybernetics, 2014, 45(9): 2013-2025.

[14] HE Qing, JIN Xin, DU Changying, et al. Clustering in extreme learning machine feature space [J]. Neurocomputing, 2014, 128: 88-95.

[15] BEZDEK J C. Pattern recognition with fuzzy objective function algorithms [M]. New York: Plenum Press, 1981: 203-239.

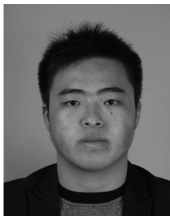
[16] HUANG Guangbin, CHEN Lei. Convex incremental extreme learning machine [J]. Neurocomputing, 2007, 70 (16/17/18): 3056-3062.

[17] HUANG Guangbin, CHEN Lei. Enhanced random search based incremental extreme learning machine [J]. Neurocomputing, 2008, 71(16/17/18): 3460-3468.

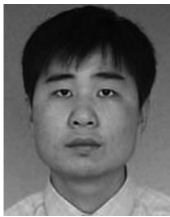
[18] DENG Zhaohong, CHOI K S, CHUNG Fulai, et al. Enhanced soft subspace clustering integrating within-cluster and between-cluster information [J]. Pattern recognition, 2010, 43(3): 767-781. .

作者简介:

刘欢,男,1993 年生,硕士研究生,主要研究方向为人工智能与模式识别、智能计算、数据挖掘。



王骏,男,1978 年生,副教授,博士,CCF 会员,主要研究方向为人工智能与模式识别、智能计算、数据挖掘。



邓赵红,男,1981 年生,教授,博士,CCF 高级会员,主要研究方向为人工智能与模式识别、智能计算、系统建模。



2017 第二届 IEEE 云计算与大数据分析国际会议

2017 the 2nd IEEE International Conference on Cloud Computing and Big Data Analysis

云计算和大数据是近年来非常热的话题,也是近年来非常重要的技术。随着大数据时代来临,在以云计算为代表的技术创新大幕的衬托下,它将在众多领域掀起变革的巨浪,大数据会逐步为人类创造更多的价值。与此同时,中央和国家今年也在力推“互联网+”以及大数据战略,以及在十三五规划中强调的创新驱动,均离不开当前信息技术中的云计算和大数据。

就目前国际国内对云计算和大数据发展趋势来看,加大对云计算和大数据的技术层面的理解,加深对云计算和大数据在推动社会发展的层面的理解,以及增强各个研究和应用领域之间的交流显得尤为重要。

为此,将于 2017 年 4 月在中国成都联合展开为期 3 天的第二届 **IEEE** 云计算与大数据分析国际会议。会议旨在促进云计算与大数据分析等领域的学术交流与合作,热忱欢迎从事相关技术研究的专家、学者和专业技术人员踊跃投稿并参加大会。

会议网站:<http://www.icccbda.com/>