



计算机科学

COMPUTER SCIENCE

基于小生境算法的空气质量模糊认知图预测

韩慧健, 刘可鑫, 林雪

引用本文

韩慧健, 刘可鑫, 林雪. 基于小生境算法的空气质量模糊认知图预测[J]. 计算机科学, 2024, 51(11A): 240300120-6.

HAN Huijian, LIU Kexin, LIN Xue. Air Quality Fuzzy Cognitive Map Forecasting Based on Niche Genetic Algorithm [J]. Computer Science, 2024, 51(11A): 240300120-6.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于多维稀疏表示的空气质量指数数据补全](#)

Data Completion of Air Quality Index Based on Multi-dimensional Sparse Representation

计算机科学, 2023, 50(8): 52-57. <https://doi.org/10.11896/jsjx.220500277>

[一种城市需水量预测的模糊认知图方法](#)

Fuzzy Cognitive Map Method for Forecasting Urban Water Demand

计算机科学, 2019, 46(11A): 47-51.

[基于白平衡修正的室外光照估计算法](#)

Outdoor Lighting Estimation Algorithm Based on White Balance Correction

计算机科学, 2019, 46(6A): 211-214.

[纹理细节保持的图像插值算法](#)

Texture Detail Preserving Image Interpolation Algorithm

计算机科学, 2019, 46(6A): 169-176.

[一种基于地形因素的空气质量空间估计方法](#)

Spatial Estimation Method of Air Quality Based on Terrain Factors LV Ming-qi LI Yi-fan CHEN Tie-ming

计算机科学, 2019, 46(1): 265-270. <https://doi.org/10.11896/j.issn.1002-137X.2019.01.041>

基于小生境算法的空气质量模糊认知图预测

韩慧健^{1,2,3} 刘可鑫^{1,2} 林雪^{1,2}

1 山东财经大学计算机科学与技术学院 济南 250014

2 山东省数字媒体技术重点实验室 济南 250014

3 山东省信息可视化与计算机工程技术研究中心 济南 250014

(hanhuijian@sdufe.edu.cn)

摘要 工业化使得全球经济取得了突飞猛进的增长,但也使得环境污染愈发严重。大气污染成为世界各国讨论的话题。文中提出一种基于小生境遗传算法的改进形式的空气质量模糊认知图预测方法。利用模糊认知图表示空气污染物以及空气质量指数的关联关系,并应用改进的小生境遗传算法优化模型,使得训练结果更接近全局最优解。文中使用2015—2021年的空气数据训练模型,在2022年的数据集上测试模型。实验结果表明,与传统的遗传算法和BP神经网络相比,所提方法预测精度更高,且泛化性能更好,证明了其有效性。

关键词 空气质量;模糊认知图;小生境遗传算法

中图分类号 TP3-05

Air Quality Fuzzy Cognitive Map Forecasting Based on Niche Genetic Algorithm

HAN Huijian^{1,2,3}, LIU Kexin^{1,2} and LIN Xue^{1,2}

1 School of Computer Science, Shandong University of Finance and Economics, Jinan 250014, China

2 Shandong Key Laboratory of Digital Media Technology, Jinan 250014, China

3 Shandong Information Visualization and Computational Economic Technology Research Center, Jinan 250014, China

Abstract Industrialization has led to the rapid growth of global economy, but it has also made environmental pollution more and more serious. Air pollution has become a worldwide hot topic around the world. In this paper, an air quality fuzzy cognitive map forecasting method based on niche genetic algorithm is proposed. This method indicates the relationship of air pollutants and air quality index by using fuzzy cognitive map, and makes the training target more close to the global best solution by using modified niche genetic algorithm. The air quality data from 2015 to 2021 is used to train the model, and the model is tested on the 2022 data. The result indicates that compared to the traditional genetic algorithm and BP neural network, the proposed method has higher prediction accuracy and better generalization performance, which proves its effectiveness.

Keywords Air quality, Fuzzy cognitive map, Niche genetic algorithm

1 引言

工业化为全球经济发展带来了巨大贡献,但也对环境造成了严重的污染,对人类健康产生了消极影响^[1]。大气污染和水污染已成为全球非常严重的问题,例如上世纪英国光化学烟雾。我国部分地区频发雾霾等问题已为全球环境敲响警钟,因此,治理环境污染刻不容缓,需要各国政府努力。自从“雾霾”这个概念广为人知后,空气质量成为广泛关注的话题。空气中的污染物包括PM_{2.5}, PM₁₀, SO₂, NO₂, O₃, CO^[2]等,各种污染物浓度决定了空气质量指数。因此,通过空气中各种污染物浓度预测空气质量具有实际的理论依据以及重要的研究价值。此外,根据预测到的空气质量,人们可以决定是否戴口罩、是否开启空气净化器等,具有广泛的现实意义。

目前,空气质量预测方法有很多种,最常见的是BP(反向传播神经网络, Back Propagation)、CRQA(交叉递归定量

分析, Cross Recurrence Quantification Analysis)等神经网络算法,但这些方法会陷入局部最优解,并且结构复杂。因此,研究者进一步改进神经网络,改进后的方法相比传统神经网络,数据的依赖性降低,同时系统更接近全局最优解。本文提出了一种基于改进小生境算法的空气质量模糊认知图预测方法,首先根据空气指数构建模糊认知图模型,以表示各种污染物浓度和空气质量指数之间的相关关系,该预测模型结构简单直观并且数据依赖性较低;此外,利用改进的小生境遗传算法优化预测模型,实现了较好的空气质量预测性能,并且在一定程度上避免了陷入局部最优解的问题,实现了预测精度的提升,从而使预测更准确。

2 相关研究

目前,预测空气质量的方法有很多种,每一种都具有各自的优势。Fang等^[3]提出时空相似LSTM(长短时记忆网络,

基金项目:国家社会科学基金(22BSH020)

This work was supported by the National Social Science Foundation of China(22BSH020).

通信作者:刘可鑫(kxliu1998@aliyun.com)

Long Short-Term Memory)模型,以北京市 2014—2018 年空气质量监测以及气象数据作为实验对象,将其与各种经典模型以及集成模型进行对比实验,误差均小于各对照组实验;Zhang 等^[4]基于 BP 神经网络预测石家庄和邢台的空气质量,发现污染物浓度越高测试结果精确度越低;Su^[5]将 BP 神经网络做了改进,提出了 LM(Levenberg-Marquardt)-BP 和 PCA(Principal Component Analysis,主成分分析法)-LM-BP,对南昌市空气质量进行了预测,结果准确率相比传统神经网络有了明显提升;Li 等^[6]使用 CRQA-DBN(深度置信网络,Deep Belief Network)-ELM(极限习算法,Extreme Learning Machine)方法,对北京奥体中心站点捕获的空气质量数据进行分析,结果显示,与 CRQA-STM(短时记忆网络,Short-Term Memory)、CRQA-DBN、CRQA-ELM、CRQA-LSTM、CRQA-CNN-LSTM 相比,CRQA-DBN-ELM 在 PM_{2.5} 浓度预测准确率上有明显优势。

但这些方法训练参数多,成本高,并且有陷入局部最优解的可能。因此,一些人提出了更好的方案,降低了数据复杂度。Han 等^[7]提出基于灰色关联分析指标筛选的 RBF(Radial Basis Function,径向基函数)神经网络的模型,相比传统神经网络,大大降低了数据的依赖;Meng 等^[8]提出了混沌遗传神经网络,相比传统神经网络更容易得到全局最优解;Li 等^[9]采用 T-S 神经网络对大同市空气质量进行预测,与 BP 神经网络相比,精确度更高。

模糊认知图(FCM,Fuzzy Cognitive Map)是认知图的一种。认知图最早由 Tolman^[10]提出,旨在能够从更加有效的视角看待心理问题。认知图有很多种,最先提出的是三值认知图。随后,三值逻辑扩展到 $[-1,1]$ 之间的模糊关系,提出了模糊认知图理论。该理论能比三值逻辑携带更多信息。FCM 可能终止于一个不动点、一个极限环,也可能终止于一个混沌奇异吸子。模糊认知图大大扩充了认知范围,也大大提高了系统研究效率,因此得到了广泛应用。本文采用基于小生境遗传算法的模糊认知图预测,模糊认知图可以降低复杂度,在数据的预测上应用广泛;而小生境遗传算法相比传统遗传算法能够让结果更接近全局最优解。因此,这两种方法组合可以很好地进行数据预测。

模糊认知图的权值矩阵确定方法有很多种,常用的有专家法^[11]、Hebbian 算法^[12]、实数编码遗传算法(RCGA,Real-coded Genetic Algorithm)、蚁群优化算法(ACO,Ant Colony Optimization)^[13]等。其中专家法是人工构建法,受不同专家的经验以及主观因素影响较大,因此预测误差较大。鉴于专家法的缺点,Hebbian 算法、RCGA 等计算算法被相继应用在权值矩阵预测中。计算算法形成的初始值是随机的,通过原始数据进行优化,以达到预测误差最小的权值矩阵。本文使用 RCGA 算法进行权值确定。

3 基于改进小生境算法的空气质量模糊认知图预测模型

3.1 数据收集

本文采用中华人民共和国生态环境部发布的北京市 2015—2022 年的空气质量数据作为数据源,各项数据如表 1

所列。由于数据太长,表 1 只列举了部分数据,其中包含了 PM_{2.5}、PM₁₀、SO₂、NO₂、O₃、CO 共 6 项空气污染物,空气质量指数作为各种污染物的结果。表中 CO 的浓度为 mg/m³,其他污染物为 μg/m³。

表 1 北京市 2015—2022 年空气质量指标
Table 1 Air quality index of Beijing from 2015 to 2022

月份	PM _{2.5}	PM ₁₀	SO ₂	NO ₂	O ₃	CO	综合空气质量指数
2015.1	98	112	35	68	62	3.8	8.02
2015.2	94	116	27	52	78	3.0	7.34
2015.3	87	146	22	55	111	2.2	7.57
2015.4	72	114	10	46	182	1.4	6.50
2015.5	56	93	9	40	242	1.2	5.89
2015.6	62	74	7	37	229	1.3	5.62
2015.7	62	68	5	36	238	1.1	5.49
2015.8	45	61	4	31	226	1.3	4.74
2015.9	50	59	6	44	180	1.4	4.94
2015.10	74	86	6	54	127	2.0	6.08
2015.11	118	132	12	61	52	4.4	8.40
2015.12	152	158	20	76	47	6.4	10.72
⋮							
2022.1	45	52	3	34	63	1.5	3.70
2022.2	22	37	3	20	77	0.8	2.39
2022.3	40	74	3	27	95	1.0	3.77
2022.4	35	80	3	22	168	1.1	4.07
2022.5	21	50	3	17	211	0.6	3.25
2022.6	23	47	3	14	236	0.9	3.43
2022.7	19	35	3	15	182	0.6	2.76
2022.8	19	34	3	13	162	0.7	2.65
2022.9	30	62	3	24	212	0.8	3.92
2022.10	40	69	3	29	112	0.8	3.80
2022.11	45	74	2	32	61	1.0	3.81
2022.12	17	41	2	24	67	1.5	2.22

用集合 A 表示 PM_{2.5}、PM₁₀、SO₂、NO₂、O₃、CO 的浓度以及空气质量指数,其中 $A = \{A_1, A_2, A_3, A_4, A_5, A_6, A_7\}$ 。由于收集数据的各项指标量纲不统一,根据式(1)对原始数据进行标准化处理。

$$A_i^* = \frac{A_i - \min(A)}{\max(A) - \min(A)} \quad (1)$$

其中, A_i^* 表示归一化后的数据。

3.2 模糊认知图的构建

不同的空气污染物浓度之间相互影响,并且它们共同决定着空气质量指数。利用模糊认知图表示 6 种空气污染物浓度以及空气质量指数之间的关联关系。模糊认知图中的节点 C 表示空气污染物的浓度以及空气质量指数, $C = \{C_1, C_2, C_3, C_4, C_5, C_6, C_7\}$ 。如图 1 所示,其中 $C_1, C_2, C_3, C_4, C_5, C_6$ 分别表示 PM_{2.5}、PM₁₀、SO₂、NO₂、O₃、CO, C_7 表示空气质量,它们的值分别用 A 进行初始化。模糊认知图的边表示各种空气污染物浓度以及空气质量指数之间的关系,边的权值用关系矩阵 W 表示,其中 ω_{ij} 表示 C_i 到 C_j 的关系,大于 0 表示正向联系,即 C_i 增加(减少)会导致 C_j 增加(减少);小于 0 表示负向联系,即 C_i 增加(减少)会导致 C_j 减少(增加),等于 0 表示无联系。 $\omega_{ii} = 0$,表示节点到其本身的权值为 0。权值矩阵的表示形式如下:

$$\begin{pmatrix} \omega_{11} & \cdots & \omega_{1n} \\ \vdots & \ddots & \vdots \\ \omega_{n1} & \cdots & \omega_{nn} \end{pmatrix} \quad (2)$$

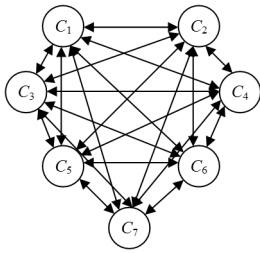


图1 空气中各项污染物的模糊认知图

Fig. 1 Fuzzy cognitive map of all pollutants in air

模糊认知图的推理公式为:

$$A_i(t+1) = f(\sum_{j \neq i} A_j(t)w_{ji}) \quad (3)$$

其中, $A_i(t+1)$ 为节点 C_i 在 $t+1$ 时刻的状态值, $A_j(t)$ 为节点 C_j 在 t 时刻的状态值, f 为阈值函数^[14]。FCM 由初始条件开始, 然后通过阈值函数传播, 再通过整个网络概念的相互作用模拟系统的动态行为。简单模糊认知图终止于稳定点或者极限环, 复杂模糊认知图可能终止于一个非稳定状态或“混沌”吸引子。设模糊认知图的所有节点为一个向量 $\mathbf{A} = (a_1, a_2, \dots, a_n)$, 则推理公式可表示为:

$$\mathbf{A}_i(t+1) = f(\mathbf{A}(t)\mathbf{W}_i) \quad (4)$$

其中, \mathbf{W}_i 是权值矩阵的第 i 列。

整个认知图迭代公式可表示为:

$$\mathbf{A}(t+1) = f(\mathbf{A}(t)\mathbf{W}) \quad (5)$$

传统的模糊认知图是单层结构, 即推理一次后就得到 $t+1$ 时刻的值。由于仅进行一次推理迭代, 因此会对复杂系统产生不利影响^[15]。鉴于传统认知图的缺陷, 本文采用多层模糊认知图, 如图 2 所示。

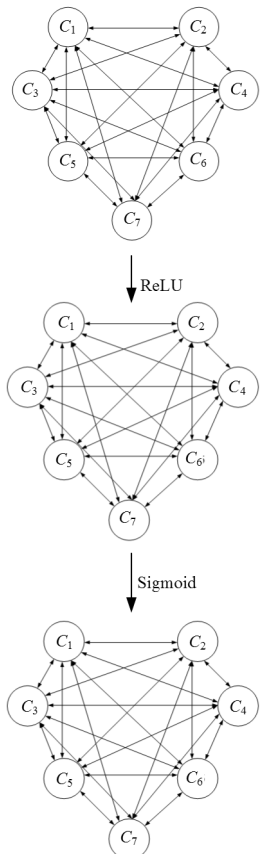


图2 2层认知图

Fig. 2 2-layer cognitive map

首先, 设置认知图层数 L , 将 t 时刻的值输入认知图, 进行计算和归一化迭代后, 将输出结果作为下一层的输入, 以此类推。当所有层都迭代完成后, 就得到 $t+1$ 时刻的最终值。本文中 $L=2$, 第一层使用 ReLU 阈值函数, 第二层使用 Sigmoid 函数。其中 ReLU 表达式如式(6)所示:

$$f(x) = \begin{cases} ax, & x \leq 0 \\ x, & x > 0 \end{cases} \quad (6)$$

本文中 $a=0.01$ 。

第二层使用 Sigmoid 函数, 表达式如式(7)所示:

$$f(x) = \frac{1}{1 + e^{-\lambda x}} \quad (7)$$

本文中 $\lambda=5$ 。

3.3 基于小生境算法的模糊认知图优化

模糊认知图的权值矩阵优化有很多种方法, 其中最常用的就是遗传算法。遗传算法是一种利用生物进化性质进行的进化算法^[16], 这种算法利用“适者生存”的原理^[17], 首先随机生成一个种群, 然后对种群的每一个个体进行编码, 编码过程中首先需要对每个个体进行适应度计算, 经过计算后利用轮盘赌法选出对应的个体, 并将其放入一个新的种群中, 再对新种群的个体进行编码、适应度计算, 以此类推, 直到达到迭代次数为止。随机种群生成后, 以后每次迭代的个体平均适应度会越来越高。由 Michalwicz 等^[18]提出的 RCGA 是一种通过实数编码^[19]和适应度函数进行的遗传算法, 这种算法的操作可以直接在解的表现型上进行^[20], 因此适用范围广泛。本文将每个节点之间的权值作为每个个体的基因, 每个个体的基因组成如式(8)所示:

$$\mathbf{W} = (\omega_{11}, \omega_{12}, \dots, \omega_{i,j+1}, \dots, \omega_{77}) \quad (8)$$

误差作适应度很好地反映了个体能否正确进行样本识别, 这样在进行一次次迭代后, 适应度高的个体就能保留下来, 而适应度低的个体就会被淘汰。本文中适应度以真实值和测量值的误差为基准, 本文只需要计算综合空气质量指数的误差, 如式(9)、式(10)所示:

$$E = \frac{1}{T} \sum_{t=1}^T |A^{(t)} - \hat{A}^{(t)}| \quad (9)$$

$$fit = \frac{1}{1+E} \quad (10)$$

其中, T 为时间序列的长度, $A^{(t)}$ 和 $\hat{A}^{(t)}$ 分别代表综合空气质量指数的真实值(表 1 中的值)和预测值, E 为测量的平均误差, fit 为个体的适应度。遗传算法流程如图 3 所示。

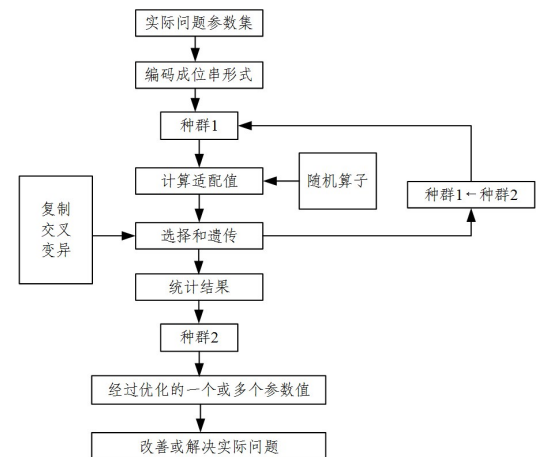


图3 遗传算法的流程

Fig. 3 Process of genetic algorithm

模型训练的具体步骤如下:

1) 首先生成一个个体数为 100 的种群, 因为模型共有 7 个节点, 所以种群内共有 49 个元素, 即 $W_i = [\omega_{i11}, \omega_{i12}, \dots, \omega_{i77}]$. 其中 W_i 是种群的第 i 个个体, ω_{ijk} 是个体 W_i 中 C_j 到 C_k 的权值. 个体的每个基因都是 $[0, 1]$ 之间的浮点值. 生成时个体内的每个基因随机产生, 若绝对值小于 0.1 则置为 0.

2) 按照式(9)、式(10)计算每个个体的适应度.

3) 计算完个体适应度后, 进行精英选择, 即选出适应度最高的 2 个个体直接放入下一代.

4) 选出除精英外的 98 个个体, 按照轮盘赌法随机选出两个个体作为父代, 其中只将一个个体放入下一代. 然后按照交叉概率进行单点交叉, 交叉时, 随机选出一个交叉点, 将父代两个个体位于交叉点后的基因对调. 在本实验中, 假设交叉点坐标为 (i, j) , 则选出的父代分别为:

$$W_1 = (\omega_{111}, \omega_{112}, \dots, \omega_{1,i,j+1}, \dots, \omega_{177})$$

$$W_2 = (\omega_{211}, \omega_{212}, \dots, \omega_{2,i,j+1}, \dots, \omega_{277})$$

经过交叉后变为:

$$W_1 = (\omega_{111}, \omega_{112}, \dots, \omega_{2,i,j+1}, \dots, \omega_{177})$$

$$W_2 = (\omega_{211}, \omega_{212}, \dots, \omega_{1,i,j+1}, \dots, \omega_{177})$$

W_1 成为待进入下一代的个体. 在完成交叉后 W_1 将按照一定的概率变异. 本文采用单点变异和多点变异. 当上一代平均适应度大于 0.9 时, 使用单点变异, 否则使用多点变异. 在本实验中, 变异的步长 MP 为 0.05, 变异值为:

$$W_{\text{new}} = W + (r - 0.5)MP \quad (11)$$

其中, W 为个体的任意一个基因, r 为 $[0, 1]$ 之间的任意取值, MP 为最大变异步长. 如果是单点变异, 则随机选取一个基因(节点到它自身的基因除外), 按照变异率进行变异; 如果是多点变异, 则该个体除节点到自身外的所有基因都按照变异率进行变异. 其中, 变异后的绝对值若大于 1, 则恢复到变异前的值. 生成一个后代后, 继续从上一代种群中选出 2 个个体遗传、交叉、变异, 直到后代种群数量达到 100 为止.

5) 将生成的后代种群继续按照步骤 2)–4) 迭代, 若迭代次数达到设定的值, 则停止迭代, 输出最终结果.

在迭代的过程中, 种群留下来的都是适应度较高的个体, 而进行交叉、变异后会产生适应度更高的个体, 种群的平均适应度和最大适应度都会提高. 传统遗传算法可以较快地得到解, 但它可能会因某些个体变异过短而出现近亲繁殖, 从而过早收敛陷入局部最优解, 因此得不到最好结果^[21-23]. 因此, 可以对遗传算法进行适当优化, 优化后可以降低陷入局部最优解的概率, 同时提高计算准确性. 目前遗传算法的优化主要是小生境算法. 小生境利用自然界中“每个物种都有特定的生存环境”的原理, 将物种的适应度差距进一步拉大, 从而增加了种群多样性. 小生境本意是特定环境的角色, 在遗传算法中, 对经过遗传、交叉、变异后的种群进行一系列处理, 使得特定距离内仅有一个个体^[24], 从而实现一个个小生境, 让适应度更高的个体更容易遗传到下一代. 小生境算法改善了传统遗传算法因过早收敛而陷入局部最优解的问题, 故常用于高维空间^[25].

在文献[25]中, 适应度函数为:

$$f_{\text{new}} = \frac{f}{2} \quad (12)$$

本文对小生境遗传算法做了改进. 因为影响适应度的是真实值和预测值的绝对误差, 而对绝对误差进行强制改动就可以改变其遗传到下一代的概率. 本文采用一种新的惩罚方法, 惩罚函数具体公式如式(13)–式(15)所示:

$$e = \frac{1}{f} - 1 \quad (13)$$

$$e_{\text{new}} = 2e \quad (14)$$

$$f_{\text{new}} = \frac{1}{1 + e_{\text{new}}} \quad (15)$$

图 4 中横坐标表示归一化后预测值和真实值的差距, 红线表示惩罚前的适应度, 蓝线表示经文献[24]中的惩罚函数惩罚后的适应度, 绿线表示经本文提出的函数惩罚后的适应度.

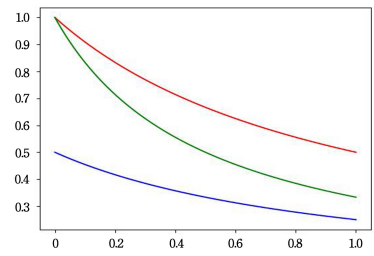


图 4 不同惩罚函数与惩罚前适应度对比

Fig. 4 Comparison of different penalty functions and pre-punishment fitness

可以看出, 与文献[24]中的函数相比, 本文提出的惩罚函数对训练误差较低(适应度较高)的个体的惩罚力度更小, 而对训练误差较高(适应度较低)的个体惩罚力度更大. 因此, 可以增大不同个体之间的适应度差距, 从而更好地实现物种多样性.

在训练开始时, 和传统遗传算法一样, 首先进行适应度计算. 适应度计算完成后的步骤如下:

1) 进行精英选择, 选出适应度最高的 2 个个体.

2) 将种群中的个体执行遗传算法的程序, 生成 100 个个体组成下一代种群, 交叉率、变异率和 3.3 节训练相同.

3) 将第 1 步选出的 2 个个体和第 2 步选出的 100 个个体合成一个新种群, 将这 102 个个体进行适应度两两比对. 按照式(16)计算两者之间的海明距离^[26], 当两者距离小于 3 时, 将适应度低的个体适应度按照式(13)–式(15)作惩罚:

$$\|x_i - x_j\| = \sqrt{(x_{ik} - x_{jk})^2} \quad (16)$$

其中, e 和 f 分别是惩罚前的绝对误差和适应度, e_{new} 和 f_{new} 分别是惩罚后的绝对误差和适应度.

4) 将个体以重新计算好的适应度选出前 100 个个体, 如果达到设定的迭代次数, 则计算完成, 输出结果; 否则重新计算适应度, 再次进行步骤 1)–3).

4 实验分析

本文使用表 1 中 2015–2021 年的空气质量数据作为训练集, 2022 年的数据作为测试集. 本文方法分别与传统的遗传算法、BP 神经网络方法进行了对比.

使用 BP 神经网络进行测试. 神经网络共有 1 个连接层, 连接层共有 30 个节点, 学习率为 0.02, 进行 10 万次迭代后, 得出的结果如表 2 所列.

表2 BP神经网络测试的结果

Table 2 Test results of BP neural network

月份	预测值	真实值	误差/%
1	3.82	3.70	3.24
2	2.53	2.39	5.86
3	3.65	3.77	-3.18
4	3.62	4.07	-11.06
5	2.66	3.25	-18.15
6	2.81	3.43	-18.08
7	2.45	2.76	-11.23
8	2.43	2.65	-8.30
9	3.31	3.92	-15.56
10	3.60	3.80	-5.26
11	3.79	3.81	-0.52
12	2.63	2.22	18.47
平均误差			8.45

在使用传统模糊认知图和小生境算法时,交叉率为0.8,变异率为0.06,最大变异步长为0.05,迭代次数为1000次。在3.3节的传统遗传算法训练结果中,在进行1000次迭代后,整个个体的最大适应度为0.9340。在测试集中,测试结

果的误差大于训练集误差。使用适应度最大的个体进行测试,得到的结果如表3所列。

表3 传统遗传算法的测试结果

Table 3 Test results of traditional genetic algorithm

月份	预测值	真实值	误差/%
1	2.72	3.70	-26.49
2	2.39	2.39	0.00
3	3.61	3.77	-4.24
4	3.72	4.07	-8.60
5	3.38	3.25	4.00
6	3.44	3.43	0.29
7	2.80	2.76	1.45
8	2.70	2.65	1.89
9	3.73	3.92	-4.85
10	3.66	3.80	-3.68
11	3.57	3.81	-6.30
12	2.31	2.22	4.05
平均误差			5.49

对应的权值矩阵如下:

$$\begin{pmatrix} 0 & 0 & -0.49 & -0.30 & 0.45 & 0.65 & 0.20 \\ 0.17 & 0 & 0 & 0.23 & -0.10 & 0 & 0.46 \\ -0.23 & 0.20 & 0 & 0.92 & -0.27 & -0.17 & 0.45 \\ 0 & 0.28 & 0.22 & 0 & 0.83 & 0.48 & -0.57 \\ 0.45 & 0.43 & -0.85 & -0.17 & 0 & 0.72 & -1.00 \\ 0.31 & -0.94 & 0 & -0.15 & 0 & 0 & -0.99 \\ 0.51 & 0.96 & 0.90 & 0 & -1.00 & -1.00 & 0 \end{pmatrix}$$

使用小生境遗传算法训练,进行100次迭代后,整个种群最大适应度为0.9321。使用适应度最大的个体测试,测试结果如表4所列。

表4 小生境遗传算法的测试结果

Table 4 Test results of niche genetic algorithm

月份	预测值	实际值	误差/%
1	3.44	3.70	-7.03
2	2.31	2.39	-3.35
3	3.93	3.77	4.24
4	4.04	4.07	-0.74
5	3.81	3.25	17.23
6	3.88	3.43	13.12
7	2.68	2.76	-2.90
8	2.64	2.65	-0.38
9	3.99	3.92	1.79
10	3.88	3.80	2.11
11	3.82	3.81	0.00
12	2.22	2.22	0.00
平均误差			4.41

对比表2—表4可以得出,BP神经网络的平均误差为8.45%,最大误差为18.47%,只有3个月的误差低于5%;传统模糊认知图的平均误差为5.49%,最大误差为26.49%,有9个月的误差低于5%;而小生境认知图的平均误差为4.41%,最大误差为17.23%,并且误差低于5%的月份有9个。可以看出,相比传统神经网络,模糊认知图在测试准确性上有了提升,而小生境遗传算法则使得准确性得到了进一步提升。传统认知图的最终适应度达到了0.9340,而小生境算法的最终适应度只有0.9321。小生境遗传算法虽然训练集适应度不如传统模糊认知图,但测试误差更小,这说明小生境遗传算法的泛化能力更强,因此取得的解更接近全局最优解。

对应的权值矩阵如下:

$$\begin{pmatrix} 0 & -0.54 & 0 & -0.98 & 0 & 0 & -0.96 \\ 0.89 & 0 & -0.90 & -0.39 & 0.80 & 0.96 & -0.77 \\ 0 & 0.40 & 0 & 0.12 & 0 & -0.67 & -0.70 \\ 0.66 & 0 & 0.37 & 0 & -0.48 & -0.81 & 0.72 \\ 0.36 & 0.33 & 0.78 & 0 & 0 & -0.86 & 0.86 \\ 0.15 & 0.87 & -0.78 & -0.70 & 0 & 0 & 0.75 \\ -0.98 & -0.97 & -0.97 & -0.49 & 0.95 & -0.73 & 0 \end{pmatrix}$$

结束语 模糊认知图在系统认知和数据预测上已取得了巨大的进步,利用模糊认知图可以进行很多科学研究,而模糊认知图的创建又有很多种方法。遗传算法是比较常用的一种方法,但缺陷是容易陷入局部最优解。小生境遗传算法可以在一定程度上弥补该缺陷。

传统遗传算法,但测试结果与真实值差距更小,这说明小生境算法能够在较低适应度下取得更精确的解,也就是说,小生境算法得到的解更接近全局最优解。因此,小生境遗传算法是遗传算法的一种理想优化,是一种增强遗传算法准确性的途径,同时也是一种构建模糊认知图权值矩阵的有效方法。

在实验中,小生境遗传算法在训练集上的适应度低于

参考文献

- [1] LI G P, QIU Z B, MIAO J Q, et al. Air Quality Prediction Model Based on the LSTM[J]. Journal of Southwest Minzu University (Natural Science Edition), 2023, 49(1): 67-73.
- [2] WANG C. Study on the Relationship Between Air Quality in Beijing and Meteorological Conditions and Its Prediction[D]. Lanzhou: Lanzhou University, 2018.
- [3] FANG W, ZHU R S. Air Quality Prediction Model Based on Spatial-temporal Similarity LSTM[J]. Application Research of Computers, 2021, 38(9): 2640-2645.
- [4] ZHANG J, WANG S G, DU L L, et al. Study on Prediction of Air Quality in Central and Southern Hebei Based on BP Neural Network[J]. Acta Agriculturae Jiangxi, 2019, 31(5): 96-102.
- [5] SU H. Research on Air Quality Prediction Based on Improved BP Neural Network [D]. Nanchang: Nanchang University, 2020.
- [6] LI Z G, QIN L L, FU D M. Air Quality Data Prediction Method Based on CRQA-DBN-ELM[J]. Electronic Measurement Technology, 2022, 45(19): 76-82.
- [7] HAN X LI B Y, GUAN Z B. Atmospheric Quality Prediction Model Based on RBF Neural Network-Markov Chain of Grey Correlational Analysis Filter Indexes[J]. Journal of Nankai University (Natural Science Edition), 2013, 46(2): 22-27.
- [8] MENG D, FAN C J, LU X D, et al. Application of Chaotic Genetic Neural Network to the Air Quality Forecast[J]. Journal of Safety and Environment, 2014, 14(4): 246-250.
- [9] LI Q, QIAO D, LI B W. Analysis of Air Quality Prediction Model Based on T-S Fuzzy Neural Network[J]. Inner Mongolia Coal Economy, 2022(17): 142-144.
- [10] TOLMAN, EDWARD C. Cognitive Maps in Rats and Men[J]. Psychological Review, 1948, 55(4): 189.
- [11] HAN HJ, SONG XF, ZHANG H. Fuzzy Cognitive Map Method for Forecasting Urban Water Demand[J]. Computer Science, 2019, 46(S2): 47-51.
- [12] LI H, CHEN HQ, MA LY, et al. A Review of Algorithm Improvement and Application of Fuzzy Cognitive Map[J]. Journal of Nanjing University (Natural Sciences), 2016, 52(4): 746.
- [13] SHANGGUAN Q Q. Research on Approaches of Air Quality Monitoring Data Prediction Based on Fuzzy Cognitive Maps [D]. Beijing: Beijing Jiaotong University, 2020.
- [14] YANG YP, YAN HB, MA TJ. Evolutionary Analysis of Battery Electric Vehicle Based on Fuzzy Cognitive Map[J]. Journal of Systems & Management, 2018, 27(2): 359-365.
- [15] LI W Z. Research on Enhanced Fuzzy Cognitive Map Method for Bank Credit Risk Control[D]. Jinan: Shandong University of Finance and Economics, 2022.
- [16] KONG S, LIU L L, SHI Y W, et al. PID Control of Desulfurization Slurry pH Value Based on Improved Genetic Algorithm[J]. Electric Engineering, 2021(16): 1-3.
- [17] LIU P, YU H. An Improved Adaptive Genetic Algorithm[J]. Ship Electronic Engineering, 2021, 41(6): 101-104.
- [18] MICHALWICZ Z, JANIKOW C Z, KRAWCZYK J B. A Modified Genetic Algorithm for Optimal Control Problems[J]. Computers & Mathematics with Applications, 1992, 23(12): 83-94.
- [19] CHI Y X. Research on Intelligent Learning Algorithm and Application of Fuzzy Cognitive Maps[D]. Xi'an: Xidian University, 2017.
- [20] XUN H F. Improved Genetic Algorithm Based on Real Coding and Its Application Optimization of Fuzzy Logic Controller[D]. Taiyuan: Taiyuan University of Technology, 2009.
- [21] HUO C J, WANG W R, YU K, et al. Theoretical Line Loss Forecast Based on Cascade Neural Network Optimized by Niche Genetic Algorithm[J]. Science Technology and Engineering, 2021, 21(23): 9897-9906.
- [22] KONG Y F, JIA L X, ZHANG J. Voice Print Recognition Based on Wavelet Analysis and Niche Genetic Neural Network Algorithm[C]// 20th CCSSTA. 2019: 523-526.
- [23] CHEN C, LIU T, CHOU J. A Novel Crowding Genetic Algorithm and Its Applications to Manufacturing Robots[J]. IEEE Transactions on Industrial Informatics, 2014, 10(3): 1705-1716.
- [24] CUI B L, XU G J, DA L L, et al. Shallow Sea Sound Speed Profile Inversion Based on Niche Genetic Algorithm[J]. Journal of Applied Acoustics, 2021, 40(2): 279-286.
- [25] LI Q, LIU Z J, WANG W J, et al. Intelligent Optimization of Lead-bismuth Reactor Core Based on Radial Basis Function Surrogate Model and Niche Genetic Algorithm[J]. Nuclear Power Engineering, 2022, 43(6): 93-100.
- [26] ZHU X N, ZHOU G J. Improved Niching Genetic Algorithm for Distribution Network Reconfiguration [J]. Journal of Fujian Computer, 2022, 38(9): 7-12.



HAN Huijian, born in 1971, Ph.D, professor, M. S supervisor, is a member of CCF (No. 21259S). His main research interests include information management, cognitive intelligence, and visual decision.



LIU Kexin, born in 1998, postgraduate. His main research interests include system recognition and forecast.