doi:10.3969/j.issn.2095-1744.2021.03.011

基于果蝇算法优化 GRNN 的生物氧化预处理温度预测

孔颜芳,南新元,石跃飞,苏比努尔·艾依来提
(新疆大学 电气工程学院,乌鲁木齐 830047)

摘 要:为了实现对氧化槽温度的精准预测,提出用改进过的果蝇优化算法来优化广义回归神经网络(GRNN)的参数。改进的方法是在标准果蝇优化算法中加入个体极值的思想,以此提高算法跳出局部极优值的能力。最后分别采用 GRNN、FOA-GRNN、IFOA-GRNN 方法建立了氧化槽温度预测模型。研究结果表明,IFOA-GRNN 预测模型的平均绝对误差(MAE)、均方根误差(RMSE)和平均相对百分比误差(MAPE)均比其它两种预测模型的低,预测精度和泛化能力更强,采用该法可以更精确地预测氧化槽的温度变化。

关键词:生物冶金氧化预处理;温度;广义回归神经网络;改进果蝇优化算法;预测
中图分类号:TF831;TP273 **文献标志码**:A **文章编号**:2095-1744(2021)03-0069-06

Optimization of Generalized Regression Neural Network Temperature Prediction Based on Improved Drosophila Algorithm

KONG Yanfang, NAN Xinyuan, SHI Yuefei, Subinur·Gayrat
(School of Electrical Engineering, Xinjiang University, Urumqi 830047, China)

Abstract: In order to predict the oxidation tank temperature accurately, an improved Drosophila optimization algorithm is proposed to optimize the parameters of the generalized regression neural network (GRNN). The improved method is to add the idea of individual extremum into the standard Drosophila optimization algorithm, so as to improve the ability of the algorithm to jump out of the local optimal value. Finally, GRNN, FOA-GRNN and IFOA-GRNN are used to establish the temperature prediction model of oxidation tank. The results show that the mean absolute error (MAE), root mean square error (RMSE) and mean relative percentage error (MAPE) of IFOA-GRNN prediction model are all lower than those of the other two prediction models, and the prediction accuracy and generalization ability are more reliable. This method can be used to predict the temperature change of the oxygen tank more accurately.

Key words: biological metallurgy oxidation pretreatment; temperature; generalized regression neural network; improved drosophila optimization algorithm; prediction

收稿日期:2020-07-20
基金项目:国家自然科学基金资助项目(61463047)
Fund: Supported by the National Natural Science Foundation of China(61463047)
作者简介:孔颜芳(1994—),女,硕士研究生,研究方向为计算机控制。
通信作者:南新元(1967—),男,教授,主要从事流程工业系统控制与优化、智能控制系统开发、现场总线技术研究与应用。
引用格式:孔颜芳,南新元,石跃飞,等. 基于果蝇算法优化 GRNN 的生物氧化预处理温度预测[J]. 有色金属工程,2020,11(2):69-74.
KONG Yanfang, NAN Xinyuan, SHI Yuefei, et al. Optimization of Generalized Regression Neural Network Temperature Prediction Based on Improved Drosophila Algorithm[J]. Nonferrous Metals Engineering, 2021, 11(2): 69-74.

生物氧化预处理技术是20世纪80年代提出的，主要利用驯化后的细菌与包裹黄金的难处理矿石如黄铁矿、硫磺铁矿、磁黄铁矿等矿石发生氧化反应，使金裸露出来[1]。在实际生产过程中，温度是氧化菌生存、繁殖的基本条件，在很大程度上影响氧化提金率，是生物氧化预处理过程中的重要参数，因此对温度的准确控制是该技术成功应用的关键之一。由于高寒高海拔地区气压低、空气密度小、昼夜温差大，氧化槽内温度控制十分困难。为此，国外诸多学者对氧化槽温度控制进行了大量研究，如南非金科(Gencor)[2]使用蛇形管水换热法对矿浆温度进行调整，使其维持在40 ℃以保证细菌活性；MEGURU等[3]和KURIYAMA等[4]分别使用不同的方法对搅拌槽的传热行为进行了仿真分析。国内更有诸多学者对氧化槽温度预测进行了研究，如梁甜等[5]将在线最小二乘支持向量回归机(OLS-SVR)用于温度的预测，杨梅等[6]将核极端学习机(KELM)应用于温度的预测。本文作者参考文献[6]对鲸鱼优化算法进行改进，来优化核极端学习机(KELM)的参数，并取得了较好的预测结果，然而改进的优化算法依旧有寻优速度慢、算法计算量大等缺陷。基于以上研究，本文提出采用改进的果蝇优化算法来优化广义回归神经网络模型的宽度系数σ，来预测氧化槽温度。相比较于杨梅等的预测结果，本文模型的平均绝对误差(MAE)、均方根误差(RMSE)和平均相对百分比误差(MAPE)均较低，预测精度更高。

基于神经网络的非线性映射能力，学者们取得了较好的成果。然而大多数神经网络的调节参数会影响它们自身的非线性映射能力，比如支持向量机需要设置调节核参数和罚参数、部分神经网络的神经元数量、输入矩阵等的设置，都会对预测模型有较大的影响，且调节参数越多，在预测模型参数调试阶段花费的时间就会越多，不利于预测的实时性。而广义回归神经网络结构简单、非线性映射能力强、设置参数单一、更有很强的容错性和鲁棒性，适宜处理非线性问题，因而被大家广泛应用。如孔雪卉等[7]将广义回归神经网络(General regression neural network，GRNN)运用到变电站设备温度预测上；皮骏等[8]将广义回归神经网络应用到航空发动机排气温度(EGT)预测中。

果蝇优化算法(FOA算法)是进化算法的一种[9-10]，是由台湾学者潘文超于2011年提出的，相比较于其他算法，该算法原理易懂、程序结构简单易懂且收敛速度较快，因此被大家广泛应用[11-14]。但是，经典的果蝇优化算法每次重复寻优过程都只是希望找到全局最优值，非常容易陷入局部极值，从而降低了算法的收敛精度。近年来不少学者都对FOA做出改进并结合GRNN应用到不同领域。本文参考文献[12]的思想，对经典的果蝇优化算法进行相应改进，结合GRNN对氧化槽温度进行预测，并得到较好的结果。

1 基于改进FOA的GRNN预测模型

1.1 广义回归神经网络(GRNN)模型的体系结构

1991年，The Lockheed Palo Alto研究实验室的Donald Specht提出了广义回归神经网络模型。如图1所示，该模型通常建立在数理统计基础上，非线性映射的能力及学习速度都很强，一般包括输入层、模式层、求和层和输出层。在学习样本中，输入层的神经元数和输入向量的维数相等，各神经元是简单的分布单元，将输入变量直接传递给模式层。模式层的神经元数与学习样本数n相等，并且各个神经元对应的学习样本不同。模式层中的神经元所对应的传递函数为：

$$P_i = \exp\left[-\frac{(X-X_i)^T(X-X_i)}{2\sigma^2}\right], i = 1,2,\cdots n \tag{1}$$

式中，P_i—模式层中每个神经元的输出；X—网络的输入变量；X_i—模式层中的各神经元对应的学习样本；σ—宽度系数。

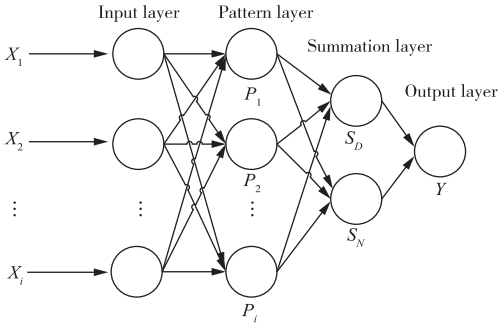


图1 GRNN网络结构图

Fig. 1 GRNN network structure diagram

GRNN的求和层存在两种神经元，对它们进行求和。第一种类型是对模式层中所有神经元输出来算术求和，并且模式层中的每个神经元与该神经元之间的连接权值是1，对应的传递函数为：

$$S_D = \sum_{i=1}^{n} P_i \tag{2}$$

第二种类型是对模式层中所有神经元的输出来加权求和，模式层神经元的输出值y_i是它的连接权

值,对应的传递函数:

$$S_N = \sum_{i=1}^{n} y_i P_i \tag{3}$$

输出层 Y 可以式(4)获得:

$$Y = \frac{S_N}{S_D} \tag{4}$$

1.2　广义回归神经网络(GRNN)理论基础

GRNN 在分类和回归预测领域有着广泛应用。该算法不易陷入局部极小值,需要优化调整的参数较少,可用于分析大规模、不稳定的数据集。因此,与其他神经网络如 BPNN 和 SVM 相比,GRNN 具有许多优点。GRNN 的结构由四部分组成:输入层用向量变量 X 馈送,模式层对输入层的输出进行非线性变换,求和层用于计算模式层输出的和,输出层产生预测结果[15]。

在回归分析期间,预测输出用:

$$\hat{y} = E[y/X] = \frac{\int_{-\infty}^{\infty} y f(X,y)\,\mathrm{d}y}{\int_{-\infty}^{\infty} f(X,y)\,\mathrm{d}y} \tag{5}$$

式中,X—自变量;y—因变量;$f(X,y)$—X 和 y 两个变量的联合概率密度函数,可以依照 Parzen-window 非参数估计方法,用随机变量 X_i 和 y_i 来估计,方法如下:

$$\hat{f}(X,y) = \frac{1}{(2\pi)^{(P+1)/2}\sigma^{(P+1)}} \times \frac{1}{n}\sum_{i=1}^{n} \exp\left[-\frac{(X-X_i)^T(X-X_i)}{2\sigma^2}\right] \times \exp\left[-\frac{(y-y_j)^2}{2\sigma^2}\right] \tag{6}$$

式中,n—样本数;P—随机变量 X 的维数;σ—宽度系数(或平滑因子)。预测输出如式(7)所示。

$$\hat{y}(X) = \frac{\sum_{i=1}^{n} \exp\left[-\frac{D_i^2}{2\sigma^2}\right] y_i}{\sum_{i=1}^{n} \exp\left[-\frac{D_i^2}{2\sigma^2}\right]} \tag{7}$$

式中,$D_i^2 = (X-X_i)^T(X-X_i)$　　(8)

GRNN 的预测精度受宽度系数 σ 的影响,需要对其进行优化:当 σ 值足够大时,预测结果接近 y 的平均值,对训练数据的拟合能力较差;当 σ 值很小时,预测结果与训练样本中的实验 y 值接近,但对外部测试集化合物的预测精度较差;当 σ 值适当时,考虑所有训练样本的实验 y 值,并对 X 附近点对应的变量 y 赋予更大的权重,产生了良好的预测性能。

1.3　果蝇优化算法(FOA)

FOA 是模拟果蝇觅食行为而得到的一种新兴群体智能优化算法。与其他蝇类相比,果蝇有较强的嗅觉和视觉能力,并且它的嗅觉器官能灵敏地识别空气中的各种气味,在接近食物时,通过它们灵敏的视觉和群体的位置将食物团团围住之后,飞往目标[16],搜寻目标过程如图 2 所示。利用 FOA 优化广义回归神经网络宽度系数 σ,并对氧化槽温度进行预测,步骤为:

1)随机初始化果蝇群的位置(x_0,y_0);

2)随机赋予果蝇个体搜寻食物的距离与方向;

$$x_i = x_0 + \mathrm{Rand}()$$
$$y_i = y_0 + \mathrm{Rand}() \tag{9}$$

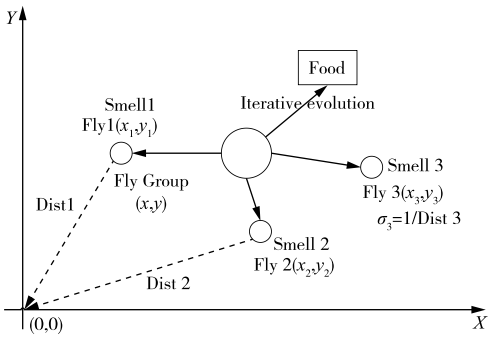


图 2　果蝇群体迭代搜寻食物示意图

Fig. 2　Schematic diagram of iterative food search in drosophila population

3)因为食物的具体位置未知,可以先确定果蝇坐标与原点的距离(D_i),再转化为待优化选择的宽度系数(σ_i),然后对 D_i 求倒数;

$$D_i = \sqrt{(x_i)^2 + (y_i)^2}$$
$$\sigma_i = 1/D_i \tag{10}$$

4)将氧化槽温度预测的均方根差(RMSE)作为味道浓度的判定函数,并计算果蝇所对应的宽度系数 σ_i 和所对应的味道浓度 $g_{RMSE}(i)$;

$$f(x) = \sqrt{\frac{\sum_{i=1}^{M} (\hat{y}_i - y_i)^2}{M}}$$
$$g_{\mathrm{RMSE}}(i) = f(\sigma_i) \tag{11}$$

式中,M—预测数量;$\hat{y}_i$—GRNN 预测的温度值;y_i—氧化槽温度的实际值;$i = 1, 2, \cdots, \mathrm{M}$;$RMSE$ 为 σ 函数的下标,表示为方均根误差,此处下标仅表示该函数计算依据,无其他特殊含义。

5)找出所有宽度系数中 g_{RMSE} 最小个体的位置:

$$[g_{\mathrm{Best}\ RMSE} \quad g_{\mathrm{Best\ index}}] = \min(g_{\mathrm{RMSE}}(i)) \tag{12}$$

其中,$g_{\mathrm{Best}\ RMSE}$、$g_{\mathrm{Best\ index}}$ 分别表示的是在此次搜寻过程中最佳宽度系数对应的味道浓度和最佳宽度系数对应味道浓度的位置。

6)保留果蝇味道浓度的最优值与其所对应的

x、y 的最佳坐标，并利用视觉飞往该位置。

$$f_{\text{Best}} = f(x)$$
$$x_{\text{Best}} = x_{\text{Best Index}}$$
$$y_{\text{Best}} = y_{\text{Best Index}} \tag{13}$$

其中，$x_{\text{Best Index}}$—搜寻过程中最佳果蝇位置的横坐标；$y_{\text{Best Index}}$—搜寻过程中最佳果蝇位置的纵坐标。

7）重复步骤 2～6，直到迭代的次数达到最大值或者满足收敛条件。

1.4 改进的果蝇(FOA)优化算法

经典果蝇算法优化广义回归神经网络宽度系数时，宽度系数值是通过对果蝇个体的坐标到原点距离求倒数得来的。但是如果距离值过大，就会造成宽度系数值很小，这样果蝇算法就可能极易陷入局部极值。为了防止果蝇优化算法陷入局部极优值，借鉴文献[12]的改进方法，通过在算法中加入个体极值的方法来调整宽度系数，有利于果蝇算法跳出局部极优值，提高算法收敛精度。改进方法为：

$$D_i = \sqrt{(x_i)^2 + (y_i)^2}$$
$$D_{\max} = \max[D(i)]$$
$$D_{\min} = \min[D(i)]$$
$$\sigma(i) = \frac{1}{D_{\max}\exp\left[\lg\left(\frac{D_{\min}}{D_{\max}}\right)\frac{n}{m}\right]} \tag{14}$$

式中，n—当前迭代次数；m—最大迭代次数。需要说明的是，当样本数量较大时，建议将样本数据集分为两部分进行训练，避免算法训练过度，影响算法的推广。

2 氧化槽温度预测建模仿真

以新疆某金矿的实测数据为实验样本，并利用 Matlab 2017b 编程软件对该方案进行代码编写与仿真分析，验证所提出的生物氧化预处理过程中温度预测方法的可行性，其算法流程图如图 3 所示。

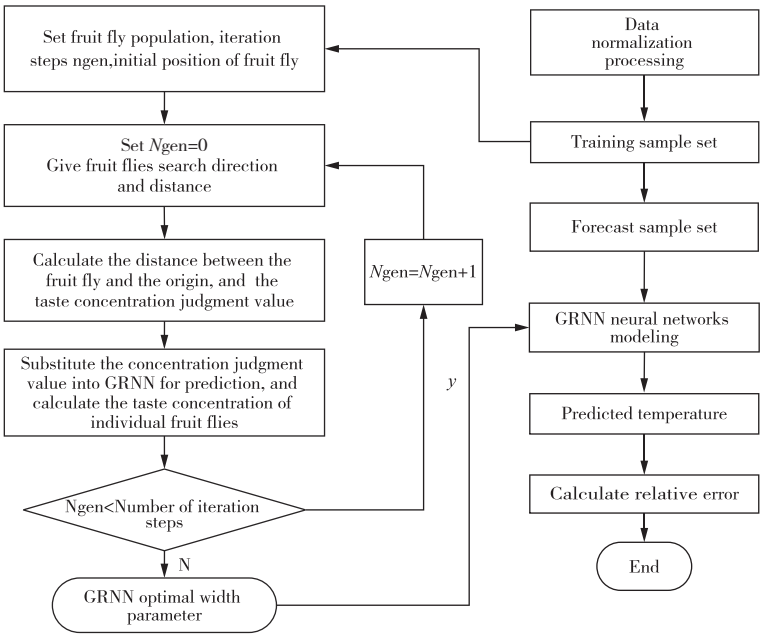


图 3 IFOA-GRNN 的预测流程图

Fig. 3 Prediction flow chart of IFOA-GRNN

2.1 实验数据采集

将 pH 值、矿浆浓度、进气量作为输入变量，温度作为输出变量，选取一级氧化槽随机一段连续时间内的 1 000 组数据作为实验样本，选择前 90%的组数据作为训练样本，后 10%的数据作为预测样本。由于实验数据都是量纲不同的变量，所以数据样本需要进行归一化处理，并将样本数据映射到[0,1]，并且训练完成后将预测模型输出数据进行反归一化操作，从而能得到实际的预测结果。

2.2 氧化槽温度预测结果

为了更好地体现改进 FOA-GRNN 模型的预测效果，运用 GRNN、FOA-GRNN、IFOA-GRNN 模型对氧化槽温度进行预测，选择 MAE、RMSE 及 MAPE 这三个性能指标来作对比，以此综合评价这三种模型的预测效果，计算公式见式(15)～(17)，计算结果如图 4～6 所示，三种模型预测结果性能比较见表 1。

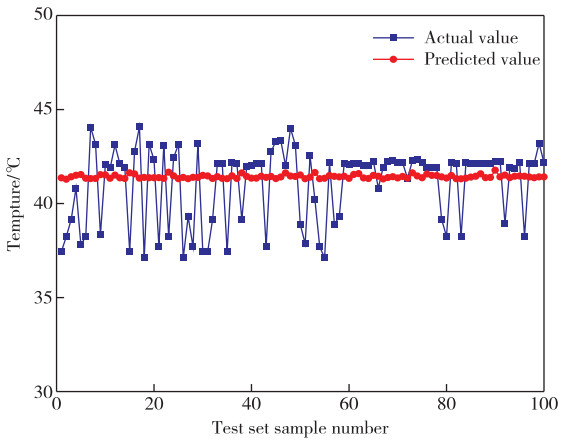


图 4　GRNN 预测值与实际值对比

Fig. 4　Comparison of GRNN predicted values and actual values

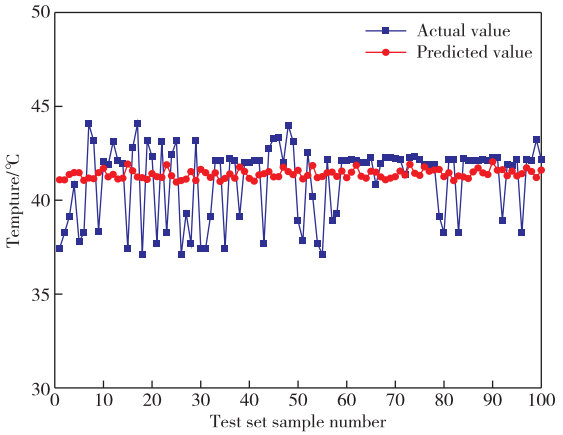


图 5　FOA-GRNN 预测值与实际值对比

Fig. 5　Comparison of predicted and actual values of FOA-GRNN

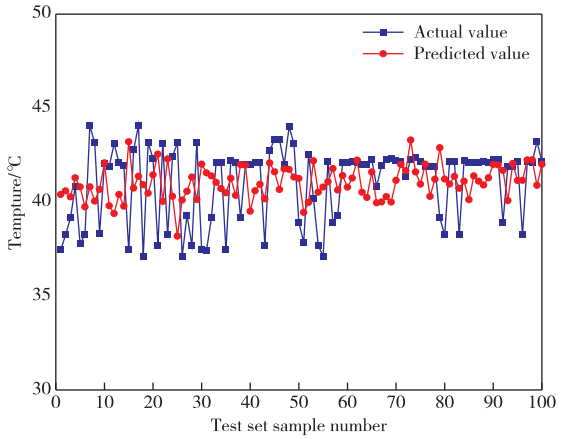


图 6　IFOA-GRNN 预测值与实际值对比

Fig. 6　Comparison of predicted and actual values of IFOA-GRNN

$$\mathrm{RMSE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i-y_i)^2} \tag{15}$$

$$\mathrm{MAE}=\frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i-y_i| \tag{16}$$

$$\mathrm{MAPE}=\frac{1}{n}\sum_{i=1}^{n}\frac{|\hat{y}_i-y_i|}{y_i} \tag{17}$$

其中,n—数据量;y_i—预测值;$\hat{y}_i$—实际值。

从表 1 中可以看出,IFOA-GRNN 模型的三个性能指标值明显比其他两个模型的低,所以该模型的预测结果明显更佳。综合比较图 4～6,可以看出,相较于其他两种模型,IFOA-GRNN 模型的预测结果与氧化槽温度的实际值更吻合,说明该模型预测精度更高。

表 1　三种模型预测结果性能对比

Table 1　Performance comparison of prediction results of the three models

Models	MAPE	RMSE	MAE
GRNN	0.006 3	0.332 7	0.227 1
FOA-GRNN	0.005 1	0.272 1	0.173 5
IFOA-GRNN	0.001 5	0.073 5	0.061 2

3　结论

1)针对生物氧化预处理复杂非线性过程,以广义回归神经网络(GRNN)为基础,采用改进的优化算法(IFOA)优化广义回归神经网络(GRNN)中的宽度系数,可以预测生物氧化预处理工艺中的氧化槽温度,为预测提供一种新的方法。

2)GRNN、FOA-GRNN、IFOA-GRNN 模型相比,IFOA-GRNN 模型的预测结果与氧化槽温度的实际值最吻合,预测精度最高。

参考文献:

[1]　杨玮.复杂难处理金精矿提取及综合回收的基础研究与应用[D].长沙:中南大学,2011,1-2.
YANG Wei. Fundamental research and application on extraction and comprehensive recovery of complex refractory gold concentrates[D]. Changsha: Central South University,2011,1-2.

[2]　NIEKERK J A V. Recent advances in the BIOX Technology[J]. Advanced Materials Research, 2009, 835:465-468.

[3]　MEGURU K, MUNEKO W, KAZUHIKO N, et al. Numerical simulation of local Heat transfer coefficients in stirred vessel with impeller for highly viscous fluids[J]. Journal of Chemical Engineering of Japan, 1999,32(1):23-30.

[4]　KURIYAMA M, OHTA M, YANAGAWA K, et al. Heat transfer and temperature distributions in an agitated tank equipped with helical ribbon impeller[J]. The Society of Chemical Engineers, Japan, 1981,

14(4):323-330.

[5] 梁甜,高丙朋.生物氧化预处理过程中氧化槽温度的预测控制[J].新疆大学学报(自然科学版),2018,35(1):108-114.
LIANG Tian, GAO Bingpeng. Predictive control of oxidation trough temperature during biological oxidation pretreatment[J]. Journal of Xinjiang University(Natural Science Edition), 2018, 35(1): 108-114.

[6] 杨梅,高丙朋.生物冶金氧化预处理过程温度预测研究[J].有色金属工程,2019,9(6):30-35.
YANG Mei,GAO Bingpeng. Research on temperature prediction of biological metallurgy oxidation pretreatment process[J]. Nonferrous Metals Engineering, 2019, 9(6):30-35.

[7] 孔雪卉,张慧芬.基于优化广义回归神经网络的变电站设备温度预测[J].中国电力,2016,49(7):54-59.
KONG Xuehui,ZHANG Huifen. Temperature forecast for substation equipment based on optimized general regression neural network[J]. Electric Power, 2016, 49(7):54-59.

[8] 皮骏,马圣,张奇奇,等.基于改进果蝇算法优化的GRNN航空发动机排气温度预测模型[J].航空动力学报,2019,34(1):8-17.
PI Jun, MA Sheng, ZHANG Qiqi, et al. Aero-engine exhaust gas temperature prediction model based on IFOA-GRNN[J]. Journal of Aerospace Power, 2019, 34(1):8-17.

[9] 潘文超.果蝇最佳化演算法:最新演化式计算技术[M].台北:沧海书局,2011.
PAN Wencao. Fruit fly optimization algorithm: a new evolutionary computing technique[M]. Taipei: Canghai Book Co., 2011.

[10] PAN W C. A new fruit fly optimization algorithm: taking the financial distress model as an example[J]. Knowledge-Based Systems, 2012, 26:69-74.

[11] 祝学昌.基于 IFOA-GRNN 的短期电力负荷预测方法研究[J].电力系统保护与控制,2020,48(9):121-127.
ZHU Xuechang. Short-term power load forecasting based on IFOA-GRNN[J]. Power System Protection and Control, 2020, 48(9):121-127.

[12] 王友卫,朱建明,凤丽洲.基于群密度的改进果蝇优化算法及在异常检测中的应用[J].工程科学与技术,2017,49(5):127-134.
WANG Youwei, ZHU Jianming, FENG Lizhou. Improved fruit fly optimization algorithm based on population density and its application in anomaly detection[J]. Advanced Engineering Sciences, 2017, 49(5):127-134.

[13] 王慧莹,吴亮红,梅盼盼,等.果蝇优化广义神经网络的风电功率短期预测[J].电子测量与仪器学报,2019,33(6):177-183.
WANG Huiying, WU Lianghong, MEI Panpan, et al. Fruit fly optimization based generalized regression neural network for wind power short-term prediction[J]. Journal of electronic measurement and instrumentation, 2019, 33(6):177-183.

[14] 王巍璋,王淳,敖鑫.基于果蝇优化算法的配电网故障定位[J].电力系统保护与控制,2019,47(18):108-114.
WANG Weizhang, WANG Chun, AO Xin. Fault location of distribution network based on drosophila optimization algorithm[J]. Journal of Electronic Measurement and Instrumentation, 2019, 47(18): 108-114.

[15] YU X L. Prediction of chemical toxicity to tetrahymena pyriformis with four-descriptor models[J]. Elsevier Inc., 2020, 190.

[16] 张健,郭星,李炜.改进果蝇优化算法在多目标搜索的应用[J].计算机工程与应用,2018,54(2):131-136.
ZHANG Jian, GUO Xing, LI Wei. Application of improved fruit fly optimization algorithm on multi object search[J]. Computer Engineering and Applications, 2018, 54(2):131-136.