

LVFOA 优化的 GRNN 在财务预警中的应用^①

赵男男¹, 王艺星², 王英博³

¹(广东海洋大学寸金学院, 湛江 524094)

²(辽宁工程技术大学 软件学院, 葫芦岛 125105)

³(辽宁工程技术大学 创新实践学院, 阜新 123000)

摘要: 针对企业财务数据复杂、非线性等特点, 提出了一种基于混沌变步长果蝇算法(LVFOA)优化广义回归神经网络(GRNN)的财务预警模型. 首先引入 Logistic 混沌映射修正 FOA 的初始值, 然后在最优初始值的基础上修正 FOA 步长为动态步长, 寻找最优 Spread 值, 最后对预测数据进行分析, 选取有代表性的指标. 改进后的果蝇算法显示了更好的全局优化和快速收敛能力, 提高了 GRNN 的预测精度. 仿真结果表明, 相对于 GRNN 模型和 FOA-GRNN 模型, LVFOA-GRNN 模型提高了预警准确率, 与财务数据的拟合度更高.

关键词: 财务预警; 果蝇算法; 广义回归神经网络; 光滑参数; 参数优化

Application of GRNN Optimized by LVFOA in the Financial Warning

ZHAO Nan-Nan¹, WANG Yi-Xing², WANG Ying-Bo³

¹(Guangdong Ocean University Cunjin College, Zhanjiang 524094, China)

²(College of Software, Liaoning Technical University, Huludao 125105, China)

³(College of Innovation and Practice, Liaoning Technical University, Fuxin 123000, China)

Abstract: Aiming at characteristics of complexity and nonlinear in enterprise financial data, this paper puts forward a financial early warning model based on General Regression Neural Network which is optimized by Logistic chaos mapping Variable step size Fruit Fly Optimization Algorithm. Firstly, the Logistic chaos mapping is used in Fruit Fly Optimization Algorithm to modify the initial value. Secondly, based on optimal initial value, we modify the step size of FOA in order to find the best Spread. Finally, we analyze the forecast data and select representative indicators. LVFOA shows better ability of global optimization and fast convergence, and it improves the prediction accuracy of GRNN. The simulation results show that the warning accuracy of new model is higher than GRNN model and FOA-GRNN model, better fitting the complex financial data.

Key words: financial warning; Fruit Fly Optimization Algorithm(FOA); General Regression Neural Network (GRNN); sooth parameter; parameter optimization

1 引言

随着我国市场经济的深入发展, 市场竞争更加激烈, 企业在寻求发展的同时, 也面临着潜在的财务危机. 财务预警系统通过分析预测以发现企业潜在的经营风险和财务风险, 并发出警告, 督促企业采取有效措施, 避免风险演变成损失. 在此背景下, 财务预警体系已经被企业广泛的应用.

国内外学者对企业财务危机的预警问题进行了各方面的研究^[1]. 预警模型的发展从上世纪 30 年代开创财务危机预警先河的单变量模型到多元线性判定模型. 1980 年, Ohlson 将四类可显著影响公司破产概率的变量衍变成多元概率比回归模型和多元逻辑回归模型, 克服了传统预警模型在假设上的局限性^[2]. 1991 年 Tam 采用了人工神经网络模型, 通过输入层、隐藏层

^① 基金项目: 2015 年广东省本科高校教学质量与教学改革工程项目(粤教高函[2015]133 号); 2015 年广东教育教学成果奖(高等教育)培育项目(粤教高函[2015]72 号); 2015 年广东省质量工程项目(粤教高函[2015]133 号); 2015 年度广东海洋大学寸金学院质量工程(CJ2015013)

收稿时间: 2016-09-30; 收到修改稿时间: 2016-11-15 [doi:10.15888/j.cnki.csa.005818]

和输出层的人工神经网络的模拟构建模型, 具有较好的模式识别能力和容错能力, 可根据数据更新进行自我学习训练, 适用于今日复杂多变的企业运作环境. 2001年吴世农、卢贤义应用 Fisher 线性判定分析、多元线性回归分析、Logistic 回归分析三种方法, 分别建立财务危机预警模型, 并得出逻辑回归模型的判定能力强.

分析可见, 对于企业财务危机预警的研究分为两种, 一是财务指标的选取, 二是财务危机预警方法的使用. 在尽量选择涵盖信息广的财务指标的同时, 应结合财务数据的特征选择有效的方法分析并处理得到的数据, 形成无限贴近非线性的高精度预警模型. 考虑到财务数据复杂和非线性的特点, 预警模型的推广应当建立在能够较好拟合和处理非线性数据的基础上. 根据果蝇算法^[2]一定的全局优化能力和广义回归神经网络隐藏层复杂的非线性网络架构, 建立了混沌变步长果蝇算法优化广义回归神经网络 LVFOA-GRNN (Logistic chaos mapping Variable step size Fruit Fly Optimization Algorithm Optimized General Regression Neural Network) 财务预警模型. 该模型一方面克服了果蝇算法因初始值选取不当导致的不稳定状况, 另一方面通过修正步长, 提高了 LVFOA 在 GRNN 网络架构中的寻优精细度, 整体提高了模型的非线性拟合能力, 较好的应用于财务预警.

2 LVFOA算法

2.1 果蝇算法

果蝇算法同蚁群算法、粒子群算法和鱼群算法等一样都是由动物的觅食行为发展而来, 属于群智能算法. 果蝇与其他物种相比感官知觉异常灵敏, 其中嗅觉与视觉尤为突出. 嗅觉器官能够很好地帮助果蝇捕

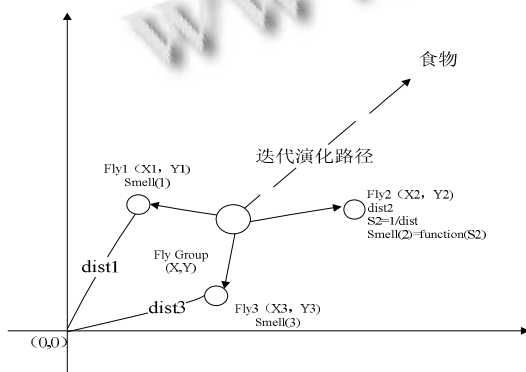


图1 果蝇寻找目标食物

捉漂浮在空气中的气味信息, 然后飞近食物位置后利用敏锐的视觉发现食物与同伴聚集的位置, 并找到目标食物. 潘文超博士通过观察此过程, 提出一种寻求全局优化的新方法, 果蝇算法^[3] (Fruit Fly Optimization Algorithm, FOA). 果蝇算法具有收敛快, 实现简单, 应用方便等优点^[4]. 果蝇寻找目标食物的过程如图 1 所示.

2.2 LVFOA

果蝇算法参数少, 改良简单, 具备较好的全局优化功能. 但其本身存在一定的缺陷. 果蝇算法初始值的选取运用了随机函数, 研究发现, 若初始值选择不当, 易陷入局部最优, 呈现不稳定的状态. 因此将果蝇算法随机产生的初始值重置为 Logistic 混沌映射^[5-7]产生的值, 使得初始值在一个看似无序实则有序的混沌环境中搜索, 并在相应混沌范围内寻找最优初始值, 进而取代原有的随机生成的初始值. 在迭代寻优的过程中, 避免陷入 GRNN 复杂网络架构的局部最优值中, 影响 FOA 在 GRNN 的寻优性能. 由于改进初始值选取后 FOA 寻找 GRNN 网络最优参数值过程缓慢, 通过测试, 虽然精度有所增加, 但适应度函数均方差未达到精度要求. 为了进一步提高收敛速率和精度, 对果蝇算法的其他参数进行优化. 增加种群规模虽可以提高收敛速度, 但同时会增加计算量并且提高程度有限. 飞行步长贯穿算法始终, 选取适当步长成为提高收敛程度的关键, 于是对果蝇算法步长进行改进^[5]. 将固定步长修正为动态步长, 经过验证当步长 h 取时收敛效果最佳. 利用逻辑混沌映射优化初始值的基础上修正步长时, 大大提高了算法收敛的精度与速度, 最终形成 LVFOA 算法^[6].

为了验证本文提出的 LVFOA 算法的性能, 采用 6 个典型的常用于优化算法性能比较的基准函数, 求解极小值, 函数形式、搜索区间、理论极值和目标精度如表 1 所示.

测试函数的维数越高、变量搜索区间越大、目标精度越高, 搜索到最优解的难度越高, 对优化算法的性能要求越高. 本文算法参数设置: 种群规模 30, 最大迭代次数 1000, 随机初始化种群位置为表一中各函数的搜索区间.

性能评估方法: 1) 固定迭代次数, 评估算法收敛的速度和精度, 并与参考文献算法进行比较; 2) 固定收敛精度, 评估算法达到目标精度所需的平均迭代次数和成功率, 并与参考文献算法进行比较; 3) 评估算法在高维、多峰函数上的性能, 并与参考文献算法进行比较;

4)评估算法的时间复杂度,并与参考文献算法进行比较.

表 1 测试函数及参数设置

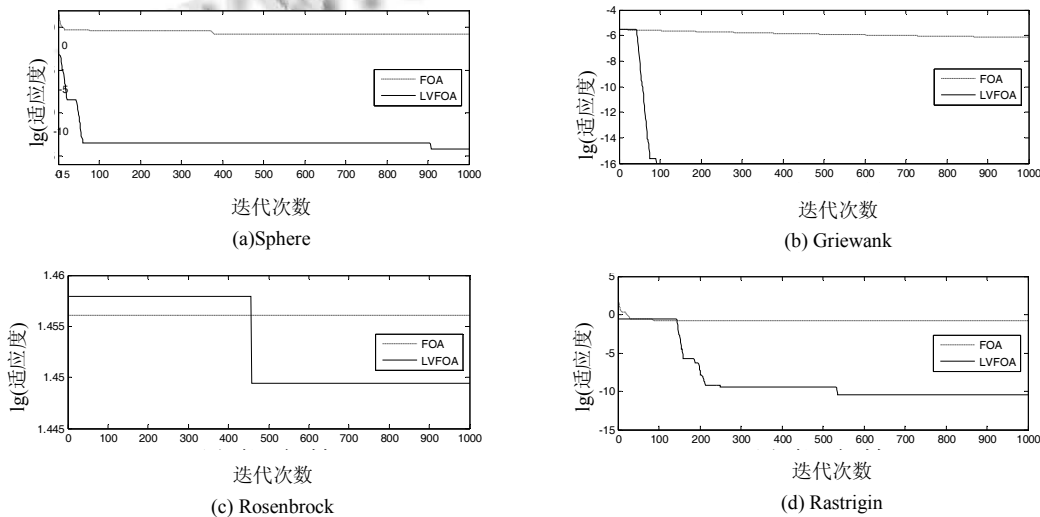
函数名	维数	公式	搜索区间	理论极值
Sphere	30	$f_1(x) = \sum_{i=1}^n x_i^2$	[-100,100]	0
Griewank	30	$f_2(x) = \frac{1}{4000} \sum_{i=1}^n x_i^2 - \prod_{i=1}^n \cos(x_i/\sqrt{i}) + 1$	[-600,600]	0
Rosenbrock k	30	$f_3(x) = \sum_{i=1}^{n-1} (100(x_{i+1} - x_i)^2 + (x_i - 1)^2)$	[-100,100]	10
Rastrigin	30	$f_4(x) = \sum_{i=1}^n (x_i^2 - 10 \cos(2\pi x_i) + 10)$	[-100,100]	0
Ackley	30	$f_5(x) = -20 \exp\left(-0.2 \sqrt{\frac{1}{30} \sum_{i=1}^n x_i^2}\right) - \exp\left(\frac{1}{30} \sum_{i=1}^n \cos 2\pi x_i\right) + 20 + e$	[-100,100]	0
Schaffer	2	$f_6(x) = \frac{\sin^2 \sqrt{x_1^2 + x_2^2} - 0.5}{(1 + 0.001(x_1^2 + x_2^2))^2} - 0.5$	[-100,100]	-1

表 2 为分别采用 FOA 和 LVFOA 两种方法^[7], 固定迭代次数为 1000 次, 分别对 6 个测试函数独立运行 100 次后的实验结果, 主要包括 100 次实验结果的最优值(Min)、最差值(Max)、平均值(Mean)和标准差(Std).

图 2 给出了 6 个测试函数的适应度对数值迭代寻优曲线(为方便显示和观察, 对函数的适应度值取以 10 为底的对数; 同时为了避免真数为 0 和纵坐标范围过大, 对函数的适应度加上 tol=1e-40 作为截止值).

表 2 FOA 与 LVFOA 算法性能比较

函数	最优值		最差值		优化均值		标准差	
	FOA	LVFOA	FOA	LVFOA	FOA	LVFOA	FOA	LVFOA
Sphere	1.70E-03	1.47E-14	7.20E-03	1.60E-08	4.20E-03	4.10E-13	1.30E-03	4.90E-04
Griewank	7.42E-07	0.00E+00	1.15E-05	3.88E-13	1.96E-06	1.94E-17	2.09E-06	2.74E-09
Rosenbrock	2.85E+01	2.81E+01	2.89E+01	2.87E+01	2.87E+01	2.87E+01	5.01E-02	4.68E-02
Rastrigin	1.52E-01	3.30E-11	1.44E+00	3.72E-07	8.20E-01	8.29E-10	2.91E-01	1.01E-08
Ackley	3.06E-02	3.01E-07	7.53E-02	3.35E-06	5.23E-02	1.18E-07	1.16E-02	1.06E-09
Schaffer	-1.00E+00	-1.00E+00	-1.00E+00	-1.00E+00	-1.00E+00	-1.00E+00	3.81E-05	2.94E-17



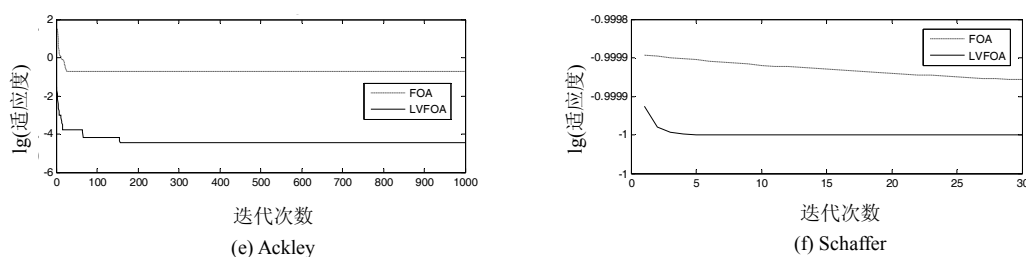


图 2 FOA 与 LVFOA 对 6 个测试函数的适应度对数值进化曲线

由表 2 可以看出, 对于单峰函数 f1 和 f3, LVFOA 的最优值、最差值、平均值和标准差都优于 FOA, 从进化曲线看, FOA 在陷入局部最优后无法跳出, 导致收敛精度不高, 而 LVFOA 可以跳出局部极值且收敛速度明显快于 FOA; 对于多峰函数 f2, f4 和 f5, LVFOA 的各项指标明显优于 FOA, 更易实现全局收敛, 100 次实验均找到函数 f2 的理论极值, 对 f4 和 f5 的最优值也分别提高了 10 个和 5 个数量级, 进化曲线显示, LVFOA 寻优速度显著优于 FOA; 强烈震荡多峰函数 Scahffer 是一个具有无数个极小值点的多峰函数, 理论极值只有一个, 在 0 处取得, 不易找到全局最优, LVFOA 在 100 次实验中都很快收敛到理论极值-1, 标准差也极小, 而 FOA 收

敛的最优值虽然接近理论极值, 但收敛速度明显不如 LVFOA, 并且寻优稳定性也略差于 LVFOA.

综合图 2 与表 2 的结果, LVFOA 比 FOA 在收敛精度, 收敛速度和跳出局部极值的能力上均有显著提高.

本文算法与参考与文献[8]中 DSFOA、文献[9]中 CFOA 和文献[10]中 IFOA 算法在相同实验条件和函数约束条件下的性能比较如表 3 所示. 实验参数设置如下: f1~f5 均为 30 维, f6 为 2 维, 对应的搜索范围分别为 ±100, ±600, ±100, ±100, ±100, ±100, sizepop=30, Maxgen=1000. 在同一实验条件下, 除 PSO 对 f1 的测试结果比 LVFOA 更优外, LVFOA 比参考文献算法具有更高的收敛精度.

表 3 LVFOA 与参考文献算法性能比较

函数	指标	LVFOA	PSO	BA	DSFOA ^[8]	CFOA ^[9]	IFOA ^[10]
Sphere	Min	1.47E-14	2.09E-22	9.62E-02	2.21E-04	1.73E-04	5.90E-11
	Max	1.60E-08	1.40E-03	2.21E-01	6.70E-03	5.40E-03	7.40E-03
	Mean	4.10E-13	9.01E-06	1.52E-01	2.80E-03	2.30E-03	1.90E-03
	Std	4.90E-04	9.95E-05	2.15E-02	1.90E-03	7.18E-04	2.20E-03
Griewank	Min	0.00E+00	3.22E-13	1.03E-02	1.04E-08	7.03E-09	0.00E+00
	Max	3.88E-13	2.60E-01	1.19E+02	1.16E-05	9.88E-06	1.44E-10
	Mean	1.94E-17	3.06E-02	1.64E+01	7.24E-07	4.76E-06	7.22E-13
	Std	2.74E-09	4.18E-02	1.71E+01	2.35E-06	1.54E-06	1.02E-11
Rosenbrock	Min	2.81E+01	1.32E+00	3.67E+01	2.85E+01	2.82E+01	2.84E+01
	Max	2.87E+01	9.92E+02	7.46E+03	2.88E+01	2.89E+01	2.88E+01
	Mean	2.87E+01	6.23E+01	6.35E+02	2.87E+01	2.87E+01	2.87E+01
	Std	3.73E-02	8.71E+01	1.40E+03	4.43E-02	6.27E-02	6.41E-02
Rastrigin	Min	3.30E-11	5.07E+01	1.03E+03	4.46E-02	6.50E-02	3.00E-10
	Max	3.72E-07	2.79E+02	5.96E+03	1.43E+00	1.36E+00	1.66E+00
	Mean	8.29E-10	1.29E+02	4.06E+03	5.90E-01	4.89E-01	3.53E-01
	Std	1.01E-08	3.87E+01	7.88E+02	3.94E-01	3.84E-01	4.40E-01
Ackley	Min	3.01E-07	8.66E-06	1.80E+01	8.40E-03	9.10E-03	5.54E-06
	Max	3.35E-06	2.00E+01	2.00E+01	7.87E-02	5.96E-02	7.21E-02
	Mean	1.18E-07	1.92E+01	2.00E+01	4.19E-02	3.94E-02	2.72E-02
	Std	1.06E-09	3.64E+00	1.87E-01	1.72E-02	7.40E-03	2.30E-02

Schaffer	Min	-1.00E+00	-9.99E-01	-5.34E-01	-1.00E+00	-1.00E+00	-1.00E+00
	Max	-1.00E+00	-9.85E-01	5.00E-01	-1.00E+00	-1.00E+00	-1.00E+00
	Mean	-1.00E+00	-9.90E-01	3.90E-01	-1.00E+00	-1.00E+00	-1.00E+00
	Std	2.94E-17	1.60E-03	1.25E-01	1.27E-06	2.77E-06	7.69E-17

LVFOA 算法比 FOA 算法^[11]显示了更好的全局优化和快速收敛能力,提高了果蝇算法在 GRNN 网络中的性能.

3 LVFOA 优化 GRNN

3.1 广义回归神经网络

广义回归神经网络 (GRNN, Generalized Regression Neural Network), 它是 RBF(Radical Basis Function, 径向基函数神经网络)的一种特殊形式^[12], 是一种基于非线性回归分析理论的前馈式神经网络模型, 主要优点在于拥有较强的非线性映射能力、较快的收敛速度、柔性的网络结构、高度的容错性和鲁棒性, 能够很好的去处理非线性的回归问题^[13]. 广义回归神经网络训练简单, 其网络链接权重值是由训练样本的输出与输入来决定, 当训练样本通过隐藏层的同时, 网络训练随即完成, 而不像前馈神经网络一样, 需要较长的训练时间和高的计算成本.

广义回归神经网络结构简单, 如图 3 所示, 由三层构成: 分别为神经元输入层(左), 径向基神经元隐含层(中)和线性神经元输出层(右). 一般只有两个隐藏层, 分别为模式层和求和层. 样本输入向量的维数作为输入层神经元数目传递给模式层, 模式层中隐藏单元的个数与训练样本的个数相同, 采用高斯函数为核函数. 求和层包含分子单元和分母单元, 分子单元将模式层各单元输出加权求和, 分母单元将模式层各单元输出求和, 最后, 输出层单元将求和层分子单元与分母单元的输出相除得到估算值^[14].

创建并训练广义回归神经网络的函数为: $net = newgrnn(P, T, Spread)$ ^[13], 其中, P 为训练数据集输入向量, T 为训练数据集输出向量, Spread 为径向基函数的光滑参数. GRNN 仿真输出为 $Y = sim(net, P)$, 其中 P 为测试数据集输入向量, Y 为测试数据集输出向量, net 为训练好的广义回归神经网络. GRNN 由 RBF(径向基函数)引申而来, 因此有且只有一个自由参数, 即 RBF 的光滑参数, 所以学习过程的目的就是去寻找最佳光滑参数 Spread 值, 本文采用均方误差

MSE 取得最优 Spread.

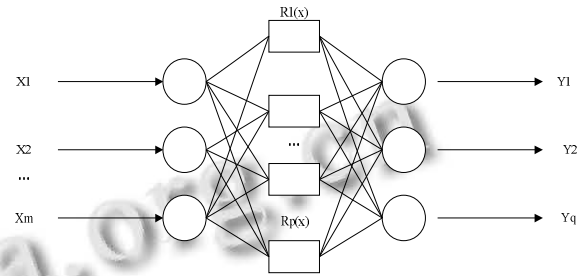


图 3 GRNN 三层神经网络结构

3.2 LVFOA 优化 GRNN

LVFOA 通过改进基本果蝇算法的初始值和步长找到最优味道浓度判定值 S_i , 代入 GRNN 的参数 Spread, 利用 Matlab 函数 newgrnn()训练得到输出值, 计算输出值与目标值的均方误差 MSE, 使之达到最小, 并保留此时的味道浓度判定值, 迭代搜索. LVFOA-GRNN 的流程图如图 4, 具体步骤如下:

Step1. 随机初始果蝇种群规模、迭代次数、以及一个二维随机数 A_axis , 第一个维数代表 x 坐标, 第二个维数代表 y 坐标.

Step2. 对 A_axis 的每个分量通过(1)式的变换映射为[0,1]范围内的混沌变量, 然后对混沌变量的每个分量进行如(3)式的混沌操作. 由于在 Logistic 映射空间[0,1]内有 3 个断点, 因此在混沌映射优化时应跳过 0.25、0.50、0.75 这三个断点.

$$Cx_i = (x_i - a_i) / (b_i - a_i) \tag{1}$$

$$x_i = Cx_i(b_i - a_i) + a_i \tag{2}$$

$$C_{x(n+1)_i} = 4C_{x(n)_i}(1 - 4C_{x(n)_i}), i = 1, 2, \dots, n \tag{3}$$

Step3. 式(1)的变量 $x_i \in [a, b]$, 可由式(1)、式(2)与混沌变量 $Cx_i \in [0, 1]$ 进行往返映射. 通过(2)式的变换, 将混沌变量映射为普通变量 A_axis_i , 带入适应度函数 $f(A_axis_i)$, 记录并选取最小值.

Step4. 重复迭代 step2 到 step3 步, 达到 $|f(A_axis_i) - f(A_axis_{i-1})| < \epsilon$, 则到一定稳定状态中止循环, 得到最佳初始值.

Step5. 给每只果蝇附上寻找目标的飞行长度与任意方向 (x_i, y_i) .

$$\begin{aligned} x_i &= X_axis + 2h \times randomvalue - h \\ y_i &= Y_axis + 2h \times randomvalue - h \end{aligned} \quad (4)$$

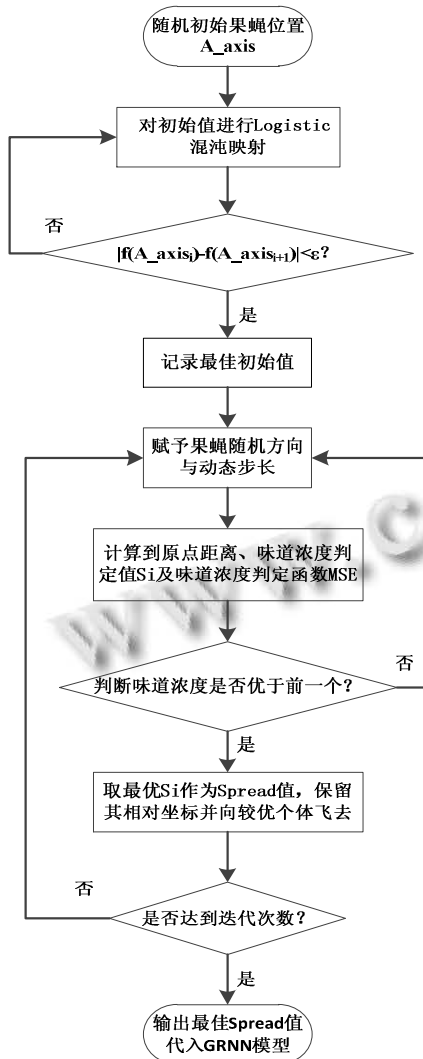


图 4 LAFOA-GRNN 流程图

h 为搜索步长, $2h \times randomvalue - h$ 为随机搜索方向. 由式(5)知 x_i 的搜索数量级在 $\pm 1/S_i$, 在 y_i 的影响下, 步长应为 $\pm k/S_i$ (k 一般取 $0.1 \sim 10$), 随机搜索方向变化范围为 $(-h, h)$, 所以步长为 k/S_i . 经过验证 k 取 0.8 时更为合适.

$$x_i = \pm \sqrt{S_i^2 - y_i^2} \quad (5)$$

Step6. 求出坐标起点到个体间距离 $Dist_i$, 再计算出味道浓度判定值 S_i , 如式(6)、(7)所示:

$$Dist_i = \sqrt{x_i^2 + y_i^2} \quad (6)$$

$$S_i = 1 / Dist_i \quad (7)$$

Step7. 将味道浓度判定值 S_i 代入 GRNN 的参数 Spread 内, 利用 Matlab 函数 newgrnn() 训练得到输出值

y_o . 并根据样本观测值计算出均方误差 $MSE(i)$ 作为味道浓度判定函数(Fitness Function), 如式(8):

$$\begin{aligned} Smell(i) &= Function(S_i) = MSE(i) \\ &= \frac{1}{n} \sum_{i=1}^n (y_o - y_p)^2 \end{aligned} \quad (8)$$

Step8. MSE 的值越小, 说明预测模型描述实验数据的精确度越好. 因此保留此时的最优味道浓度值 S_i 作为 GRNN 的 Spread 值, 并保存与其相对应的果蝇个体 (x, y) 坐标值, 此时果蝇群体利用视觉往该位置飞去.

$$\begin{aligned} [bestSmell \ bestIndex] &= \max(Smell) \\ Smell_{best} &= bestSmell \end{aligned} \quad (9)$$

$$X_axis = X(bestIndex)$$

$$Y_axis = Y(bestIndex)$$

Step9. 迭代寻优, 重复执行步骤 3 到步骤 8, 并判断味道浓度是否优于前一次迭代的味道浓度, 若是, 则执行步骤 8.

Step10. 直到达到循环次数, 将存下的最佳 Spread 值带入广义回归神经网络中实行预警得到最后的预警成果.

4 模型构建与训练

4.1 指标选取与样本分类

4.1.1 财务指标的选取

从企业偿还能力、盈利能力、运营能力、发展能力以及现金流量这五方面出发确定初始的 21 个财务指标^[15], 影响程度各不相同. 为了提高预警模型的效率和实用性, 对 21 个财务指标进行分析研究, 筛选出尽可能少且具有代表性的指标用来反映尽可能多的信息, 使筛选后的指标能有效的区分 ST 公司和非 ST 公司. 首先采用显著性检验方法 T 检验对财务指标进行第一次筛选. 使用 SPSS 统计分析软件对搜集到的 80 家训练样本公司被特别处理前两年的数据进行显著分析检验, 假设 ST 公司与非 ST 公司指标无显著差异, 结果见表 4.

表 4 各指标 T 检验显著性水平

指标名称	指标序号	T 检验显著性水平
总资产利润率	X1	0.018**
成本费用利润率	X2	0.000*
销售净利率	X3	0.21
净资产收益率	X4	0.015**
现金比率	X5	0.011**
流动比率	X6	0.018**

速动比率	X7	0.51
每股收益	X8	0.009*
利息保障倍数	X9	0.142
资产负债率	X10	0.032**
主营业务收入增长率	X11	0.065
净资产增长率	X12	0.796
总资产增长率	X13	0.001*
应收账款周转率	X14	0.074
存货周转率	X15	0.016**
固定资产周转率	X16	0.082
总资产周转率	X17	0.005*
流动资产周转率	X18	0.028**
销售现金比率	X19	0.318
经营现金负债率	X20	0.051
现金流量比率	X21	0.000*

表 4 中*和**分别代表 T 检验结果在 1%和 5%置信水平下的统计显著。从表中可以看出：现金流比率、流动资产周转率、总资产周转率、存货周转率、总资产增长率等一共 12 个指标通过了显著性检验。虽然各个指标显著性水平不同，但是为了力求全面，将这 12 个指标都入选作为第一次筛选结果，进入第二次筛选。利用 SPSS 软件做主成份分析^[16]，图 5 为主成分分析分析结果，为了更加充分地反映企业信息，选取了 91.54%的累计贡献率，这时主成份因子共有 5 个，包含了原有信息量的 91.54%，可以取代原有的 12 个指标。最终选取经营现金负债率、总资产增长率、总资产周转率、每股收益和成本费用利润率五项指标共同构成企业财务危机预警指标体系。

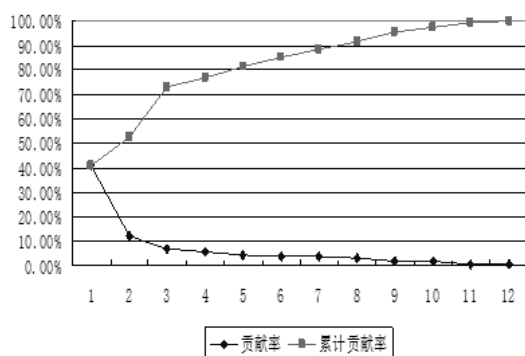


图 5 各指标主成分分析

4.1.2 样本数据分类

为了真实反映模型的应用能力，将选取的 104 家上市公司分为 5 组，其中 2010 年、2011 年首次被特别处理的 40 家 ST 公司与 40 家配对的健康公司作为训练

样本组，提取 ST 公司的 T-1 年、T-2 年数据用于模型建立与检测；另外 2012 年首次被特别处理的 12 家上市公司和 12 家配对的健康公司作为预测样本组，以便检验模型预测效果。将被特别处理的 ST 上市公司与非特别处理的健康公司分为两类，用 1、0 编码分组，其中 ST 公司类别为 1，健康公司类别为 0。部分数据如图 6 所示。

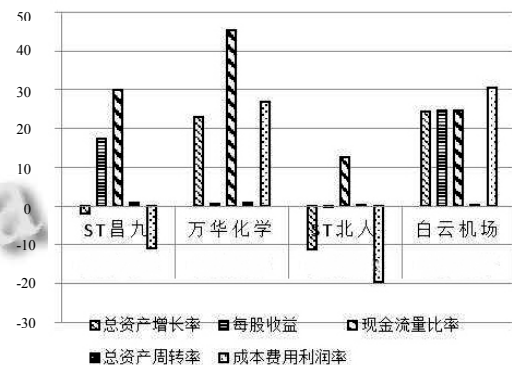
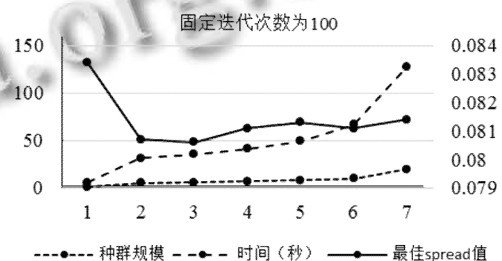


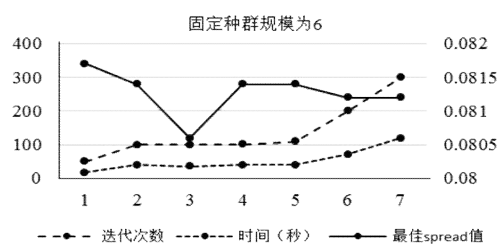
图 6 财务预警部分数据

4.2 模型参数设置

在仿真训练中，将选定的样本分为 5 组，其中 4 组数据作为训练数据，1 组作为预测数据，交叉验证，运用 LVFOA 算法寻找 GRNN 参数^[17]。为了使得 LVFOA-GRNN 模型预警效果最佳，实验中不断调整 LVFOA 的群体数量和循环次数，得到不同参数下的最佳 Spread 值，如图 7(a)和(b)所示。



(a) 固定迭代次数



(b) 固定种群规模

图 7 固定种群的迭代次数及规模

由图 7(a)可知, 固定迭代次数为 100, 在不同的种群规模下, 当种群规模为 6 时, Spread 值达到最小, 为 0.0806; 由图 7(b)可知, 当固定种群规模为 6, 变换迭代次数时, 也可以得到最佳迭代次数为 100, 当迭代次数达到 200 以上时, 最佳 Spread 值保持 0.0812 不变, 这时只会增加搜索时间. 搜索时间随着种群规模和迭代次数的增大而增大, 而均方误差除了迭代次数为 100 种群规模为 1 时收敛到 0.3411, 其他情况下均收敛到 0.3410, 但是 Spread 值却不相同, 这说明迭代次数和种群规模的选取会影响能否找到全局最优值. 当种群规模为 6, 迭代次数为 100 时果蝇针对 Spread 值寻优轨迹和均方误差收敛情景分别为图 8(a)和(b)所示. LVFOA 优化 GRNN 时, MSE 在迭代次数为 22 时趋于平稳, 在极小偏差值为 0.3410, 果蝇坐标值是(12.0749, -2.1639)时得到最优 Spread 值 0.0806, 由此建立好最佳模型.

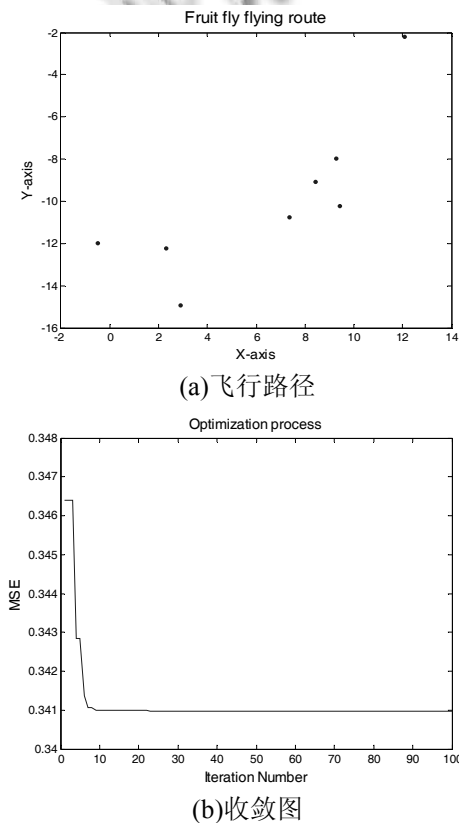


图 8 寻优轨迹和均方误差收敛情况

5 测试及结果分析

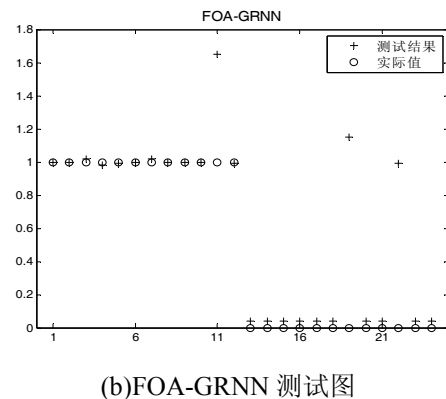
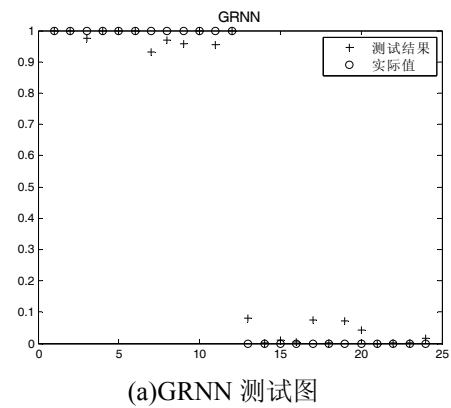
为了检验 LVFOA-GRNN 模型效果, 分别将 GRNN 模型、FOA-GRNN 模型和 LVFOA-GRNN 模型同时应

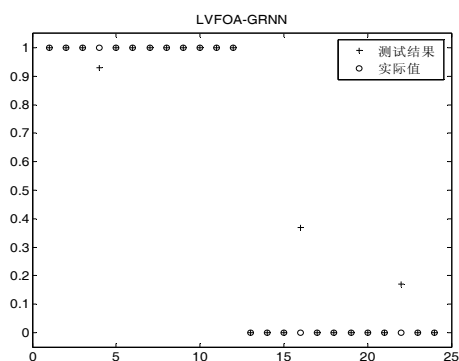
用于训练财务预警的数据 e. GRNN 网络模型采用神经网络工具箱来完成, 用交叉验证函数将数据分为五组, 经过 GRNN 训练的 Spread 值为 0.9, 三项流量预测的误差为 0.45682. FOA-GRNN 模型将选定的样本分为五组, 交叉验证, 利用 GRNN 的 Matlab 工具箱, 运用基本果蝇算法搜索 newgrnn() 的 Spread 参数为 0.124.

在允许预警结果误差小于绝对值 0.05 时, 三个模型比较得到的结果如表 5 和图 9(a)、(b)和(c)所示, GRNN 网络模型和 FOA-GRNN 模型预警准确率均低于本文模型. 改进的果蝇算法使 GRNN 的性能达到最优, 充分激发出了 GRNN 的非线性能力. LVFOA-GRNN 模型在训练过程中能够充分学习所选取的财务指标数据, 构建良好的网络架构, 使得预测准确率得到了提升.

表 5 准确率对比表

	ST 公司		非 ST 公司		合计	
	正判数	正确率	正判数	正确率	正判数	正确率
GRNN	7	58.33%	8	66.67%	15	62.50%
FOA-GRNN	9	75%	10	83.33%	19	79.16%
LVFOA-GRNN	11	91.67%	10	83.33%	21	87.5%





(c) LVFOA-GRNN 测试图

图9 模型测试图

6 结语

针对缺乏有效的企业危机预警模型,通过剖析财务预警数据的特征分析选取代表性指标,然后依次修正果蝇算法初始值和步长,测试选取和利用最优参数LVFOA寻找GRNN中的最佳Spread值,提高了GRNN的收敛速度和预测准确率,最终建立了LVFOA-GRNN预测模型。LVFOA-GRNN模型充分贴近非线性的财务数据,预警准确率基本达到了87.5%,与GRNN网络模型和FOA-GRNN模型相比具有较高的适用性。

参考文献

- 1 吴芑,吴应宇,仲伟俊.国内外企业财务预警模型比较研究.生产力研究,2007,31(9):136-138.
- 2 Ohlson JS. Financial ratios and the probabilistic prediction of bankruptcy. Journal of Accounting Research, 1980, (19): 109-131.
- 3 潘文超.果蝇最佳化演算法.台中:沧海书局,2011.
- 4 潘文超.应用果蝇优化算法优化广义回归神经网络进行企业经营绩效评估.太原理工大学学报,2011,29(4):1-5.

- 5 马超,董玲.果蝇优化算法(FOA)步长改进及其多元函数最优化方法.数学学习与研究,2013,1(13):90-92.
- 6 刘道华,原思聪,兰洋,等.混沌映射的粒子群优化方法.西安电子科技大学学报,2010,37(4):764-769.
- 7 袁小芳,王耀南,吴亮红.基于混沌变量的模式搜索法及其应用.湖南大学学报,2007,34(9):30-33.
- 8 韩俊英,刘成忠,王联合国.动态双子群协同进化果蝇优化算法.模式识别与人工智能,2013,26(11):1057-1067.
- 9 王震宇,梁雪春.基于CFOA的混合核SVM参数优化.计算机工程与设计,2016,37(3):773-777.
- 10 段艳明,肖辉辉.一种新的自适应步长果蝇优化算法.河南师范大学学报(自然科学版),2016,44(1):161-168.
- 11 程慧,刘成忠.基于混沌映射的混合果蝇优化算法.计算机工程,2013,39(5):218-221.
- 12 Specht DF. A General Regression Neural Network. Trans. on Neural Networks,1991,2(6):568-576.
- 13 刘勇,王莹,邹慧挺.改进果蝇优化算法优化广义回归神经网络的预测问题研究.世界科技研究与发展,2014,36(3):272-276.
- 14 范良,赵国忱,苏运强.果蝇算法优化的广义回归神经网络在变形监测预报中的应用.测绘通报,2013,1(11):87-89.
- 15 廖志文.基于CBR与灰色关联度的财务危机预警.计算机工程,2012,38(1):174-176.
- 16 林海明,杜子芳.主成分分析综合评价应该注意的问题.统计研究,2013,30(8):25-31.
- 17 李珺,宋文龙,朱良宽.一种新颖的关联规则数据挖掘算法.辽宁工程技术大学学报(自然科学版),2014,33(6):846-849.
- 18 翟悦,郭文书,王立娟.利用扩展概念格进行关联分类的算法.辽宁工程技术大学学报(自然科学版),2015,34(11):1280-1284.