

# 反向学习的灰狼算法优化及其在 交通流预测中的应用

张兴辉<sup>1,3</sup>, 樊秀梅<sup>1</sup>, 阿喜达<sup>1,4</sup>, 樊书嘉<sup>2</sup>, 武文瑜<sup>1</sup>

(1. 西安理工大学自动化与信息工程学院, 陕西西安 710048; 2. 清华大学计算机科学与技术系, 北京 100084;  
3. 安康学院电子与信息工程学院, 陕西安康 725000; 4. 包头师范学院信息科学与技术学院, 内蒙古包头 014030)

**摘要:** 本文针对灰狼优化(Grey Wolf Optimizer, GWO)算法平衡全局探索和局部搜索能力的不足, 提出了一种基于反向改进的灰狼算法(Opposition Learning Grey Wolf Optimizer, OLGWO), 来优化预测模型的超参数, 以提高其用于交通流预测的精度与鲁棒性. 本算法在迭代过程中采用了反向学习策略, 并引入了等级相关概念, 主要通过计算普通狼与目标狼的 Spearman 相关系数, 并根据其值来选择性地更新狼种群. 实验先对 12 个标准测试函数对比了四种算法 OLGWO、TGWO(Transformed Grey Wolf Optimizer)、GWO、PSO(Particle Swarm Optimization), 得到了寻优均值和标准差, 验证了 OLGWO 算法具有突出的性能优势; 然后采用美国加州公路交通流数据, 在不同缺失率下比较了四种算法优化的反向传播(Back Propagation, BP)网络模型, 结果显示, OLGWO-BP 模型预测精度比其它三种模型最高分别有 1.95%、3.98% 和 11.07% 的提升, 同时表现出更好的稳定性.

**关键词:** 智能交通; 交通流预测; 灰狼优化(GWO); BP神经网络; 反向学习(OL)

**中图分类号:** U491.1      **文献标识码:** A      **文章编号:** 0372-2112(2021)05-0879-08

**电子学报 URL:** <http://www.ejournal.org.cn>

**DOI:** 10.12263/DZXB.20200915

## Grey Wolf Optimization Based on Opposition Learning and Its Application in Traffic Flow Forecasting

ZHANG Xing-hui<sup>1,3</sup>, FAN Xiu-mei<sup>1</sup>, SHAN Axida<sup>1,4</sup>, FAN Shu-jia<sup>2</sup>, WU Wen-yu<sup>1</sup>

(1. Department of Automation and Information Engineering, Xi'an University of Technology, Xi'an, Shaanxi 710048, China;

2. Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China;

3. College of Electronics and Information Engineering, Ankang University, Ankang, Shaanxi 725000, China;

4. School of Information Science and Technology, Baotou Teachers' College, Baotou, Inner Mongolia 014030, China)

**Abstract:** Focusing on the weakness of the balance capability between global exploration and local search in the classical GWO algorithm, a novel grey wolf algorithm based on opposition learning (OLGWO), which can evolve the hyper-parameters of forecasting model, is proposed to improve the accuracy and enhance the robustness of traffic flow forecasting models. This algorithm is designed to take advantage of opposition learning strategy with the iterative process, and exploits the concept of rank correlation that can describe the Spearman correlation coefficients between the target wolf and the common wolves, and then selectively updates the each wolf of the whole population according to their values. Firstly, the performance comparison of four algorithms (OLGWO, TGWO, GWO, PSO), based on 12 benchmark functions, is conducted in terms of the two metrics, namely the optimization means and standard deviations. The results verify the outstanding performance of the proposed algorithm. Furthermore, based on the California highway traffic flow data, the four models optimized by the concerned algorithms are compared under different loss rates. The results show that the prediction accuracy of OLGWO-BP is higher than that of the others by 1.95%, 3.98% and 11.07%, respectively, and the stability is better.

**Key words:** intelligent transportation; traffic flow prediction; grey wolf optimizer (GWO); BP neural networks; opposition learning

## 1 引言

交通流预测可提供及时和准确的前瞻性交通流量信息,预测结果的发布,有助于调节车辆在路网中的合理分布,也可以帮助政府机构、商业部门和出行者做出更好的交通决策.由于交通流量受随机因素影响多,导致其随机性强,预测难度大,交通流量的准确实时预测仍然是一个热点和难点问题<sup>[1]</sup>.

人工神经网络由于具有拟合复杂非线性特征的能力,可以准确学习交通流的不确定性,但存在收敛速度慢,容易陷入局部极值的缺点. Fabio Moretti<sup>[2]</sup>等用统计方法优化人工神经网络,提出了一种基于神经网络的混合模型对交通流进行预测,预测结果较优,但人工神经网络的阈值为确定值,降低了预测精度. Zhao<sup>[3]</sup>等人提出了一种考虑时空特性的长短期记忆(LSTM)网络用于交通流预测,但存在隐藏层数及其神经元个数难以确定问题,根据经验设置参数对预测性能影响较大.随着新兴元启发式优化智能算法的出现,越来越多的学者开始将群智能算法与神经网络相结合,以解决神经网络预测模型精度低和调参困难的问题.蔡亮<sup>[4]</sup>等将改进粒子群算法和多层双向长短期记忆(BiLSTM)神经网络相结合用于应用负载预测,提高了预测精度. Mirjalili<sup>[5]</sup>提出一种新的元启发式算法—灰狼算法(Grey Wolf Optimizer, GWO),较其他元启发式算法具有调整参数少、便于理解、收敛速度快的优点,但存在对多峰函数求解时收敛速度慢,容易陷入局部最优解的缺点.傅蔚阳<sup>[6]</sup>等使灰狼个体移动方向满足反向学习策略,改善了算法的寻优范围,但算法的收敛速度依然较慢.张文胜<sup>[7]</sup>等提出一种改进灰狼算法(Transformed Grey Wolf Optimizer, TGWO)优化BP神经网络,对标准灰狼算法中的收敛因子和位置更新公式进行改进,对单峰函数改善了精度的同时对多峰函数的效果变差,收敛速度也有提高空间.

根据上述研究现状的分析,目前利用元启发算法进行函数值寻优的技术其平衡全局探索、局部搜索和收敛速度的能力仍不足,本文提出一种反向学习策略来进一步改进GWO算法的收敛速度和求解精度.针对误差逆向传播(Back Propagation, BP)神经网络的性能易受权值影响的问题,基于真实交通数据,利用改进后的算法OLGWO优化BP网络,并将四种PSO-BP、GWO-BP、TGWO-BP和OLGWO-BP优化模型基于两种指标对比,验证模型的有效性.

## 2 基于反向学习的灰狼算法

### 2.1 灰狼算法

在灰狼群体中,灰狼个体的社会层级常用 $\alpha$ 、 $\beta$ 、 $\delta$ 和

$\omega$ 四层表示<sup>[8]</sup>. $\alpha$ 狼为头狼,领导整个灰狼群体; $\beta$ 狼为协助者,辅助头狼决策; $\delta$ 狼主要负责侦查、放哨等任务; $\omega$ 狼处于底层,听从前三层的指挥.灰狼算法是对灰狼群等级制度和捕猎行为的数学模拟,假设种群数量为 $N$ ,搜索空间为 $D$ 维,则种群可表示为 $X = \{X_1, X_2, \dots, X_N\}$ ,第 $i$ 个灰狼个体可表示为 $X_i = \{X_i^1, X_i^2, \dots, X_i^D\}$ ,则狼群的包围行为公式表示为

$$D = |C \cdot X_p(t) - X(t)| \quad (1)$$

$$X(t+1) = X_p(t) - A \cdot D \quad (2)$$

式中: $X_p(t)$ 为种群迭代至第 $t$ 代时猎物的位置; $X(t)$ 为第 $t$ 代灰狼个体的位置; $D$ 为灰狼个体和猎物之间的距离. $A$ 和 $C$ 为系数向量,计算公式为

$$A = 2a(r_1 - 1) \quad (3)$$

$$C = 2r_2 \quad (4)$$

式中 $r_1, r_2$ 为标量在 $[0, 1]$ 之间的随机向量; $a$ 为收敛因子, $a$ 的值随着迭代次数的增加从2线性递减至0.

包围行为中,调整 $A$ 和 $C$ 的值可以实现对猎物的搜索,通过 $r_1$ 和 $r_2$ 的随机设定,可以搜索到周围的 $n$ 个方向<sup>[9]</sup>.由式(2)可知,当 $|A| > 1$ 时,下一次迭代得到的灰狼个体将从当前位置远离猎物,即候选解趋向由猎物向外发散,以进行全局搜索,搜索其他潜在最优解;当 $|A| < 1$ 时,灰狼个体的下一位置将处于当前位置和猎物之间,即候选解趋向于向猎物聚集,以进行局部搜索,逼近最优解.

灰狼将猎物包围后由最靠近猎物的 $\alpha$ 、 $\beta$ 、 $\delta$ 按照下述公式更新位置进行捕猎:

$$D_\alpha = |C_1 \cdot X_\alpha(t) - X(t)| \quad (5)$$

$$D_\beta = |C_2 \cdot X_\beta(t) - X(t)| \quad (6)$$

$$D_\delta = |C_3 \cdot X_\delta(t) - X(t)| \quad (7)$$

$$X_1(t+1) = X_\alpha(t) - A_1 \cdot D_\alpha \quad (8)$$

$$X_2(t+1) = X_\beta(t) - A_2 \cdot D_\beta \quad (9)$$

$$X_3(t+1) = X_\delta(t) - A_3 \cdot D_\delta \quad (10)$$

$$X(t+1) = \frac{X_1(t+1) + X_2(t+1) + X_3(t+1)}{3} \quad (11)$$

式中: $X_\alpha(t)$ 、 $X_\beta(t)$ 、 $X_\delta(t)$ 为种群迭代至第 $t$ 代时 $\alpha$ 、 $\beta$ 、 $\delta$ 的位置, $X(t)$ 为第 $t$ 代灰狼个体的位置, $A_1$ 和 $C_1$ 、 $A_2$ 和 $C_2$ 、 $A_3$ 和 $C_3$ 分别表示 $\alpha$ 、 $\beta$ 、 $\delta$ 的系数向量, $X_1(t+1)$ 、 $X_2(t+1)$ 、 $X_3(t+1)$ 分别表示 $\alpha$ 、 $\beta$ 、 $\delta$ 个体在 $(t+1)$ 迭代后的位置向量.

### 2.2 改进策略

前面标准灰狼算法通过式(5)~式(11)不断调整猎物方向和与猎物的距离,多次迭代后捕获猎物,此时头狼 $\alpha$ 所在的位置对应的适应度即为最优解.从迭代过程可以发现,灰狼个体随着迭代次数的增加将朝着同区域运动,使狼群的趋同性变得严重,增加了陷入局部最优的风险. Tizhoosh<sup>[10]</sup>提出了反向学习的概念,说

明了在取值区域内生产当前个体的反向解个体,综合选取较优者可提高近一半的可能性接近全局最优解.采用反向学习策略有助于改善陷入局部最优解和提升算法的寻优速度,具体改进点如下:

(1) 引入 Spearman 相关系数

本文的反向学习策略在每次迭代过程中以  $\alpha$  狼为最优目标,为了发现  $\omega$  狼中潜在的更优的狼,需要衡量每个  $\omega$  狼与  $\alpha$  狼的相关性,本文引入 Spearman 相关系数,它是利用两变量的秩次大小作线性相关分析,对原始变量的分布不作要求,属于非参数统计方法,有着对异常值敏感度低和数据不须满足正态性假设等优点. Spearman 相关系数的计算公式如式 (12) 所示,假设  $\alpha$  狼为  $(x_1^\alpha, x_2^\alpha, \dots, x_N^\alpha)$ ,  $\omega$  狼为  $(x_1^\omega, x_2^\omega, \dots, x_N^\omega)$

$$r_s = \frac{\sum_{i=1}^N (r_i - \bar{r})(s_i - \bar{s})}{\sqrt{\sum_{i=1}^N (r_i - \bar{r})^2} \sqrt{\sum_{i=1}^N (s_i - \bar{s})^2}} \quad (12)$$

式中,  $r_i$  和  $s_i$  分别表示  $\alpha$  狼和  $\omega$  狼的秩,当其中的值出现相等的时候,该值对应的秩为这几个值对应的平均值<sup>[11]</sup>. 为计算简单,可通过计算  $\alpha$  狼和  $\omega$  狼的等级差值计算,公式如式 (13) 所示:

$$r_s = 1 - \frac{6 * \sum_i (d_i)^2}{N(N^2 - 1)} \quad (13)$$

式中,  $r_s$  为等级相关系数;  $N$  为狼的维度;  $d_i = d_i^\alpha - d_i^\omega$ , 是  $\alpha$  狼和每个  $\omega$  狼对应值的等级差,  $\alpha$  狼和  $\omega$  狼按等级表示为  $(d_1^\alpha, d_2^\alpha, \dots, d_N^\alpha)$  和  $(d_1^\omega, d_2^\omega, \dots, d_N^\omega)$ .  $r_s$  的取值范围在  $[-1, 1]$  之间, 绝对值越大相关性越强, 取值符号表示相关的方向.

(2) 反向学习策略

在搜索给定函数的最优解的过程中,通常会初始随机选择一些解,如果该解接近最优解则收敛很快,如

果远离最优解,则可能收敛慢或者朝着背离最优解的方向发展. 本文提出在经典灰狼搜索迭代过程中引入反向学习策略,在每次搜索过程中把  $\alpha$  狼视为最优解,依据 Spearman 相关系数发现偏离最优解的  $\omega$  狼,通过定义反向点,从其对应的反向点中发现潜在的更优的解用于更新狼种群. 为说明反向学习策略,这里假设解的维度为 2,如图 1 所示,假设某次迭代过程中  $\alpha$  狼所处位置为图 1 中的目标位置,  $(a, b)$  属于  $\omega$  狼,用两点之间的直线距离表示其适应度,  $(a, b')$ 、 $(a', b)$  和  $(a', b')$  是  $\omega$  狼的反向点,图中明显看出  $d_2 < d_3 < d_1$ , 因此  $(a, b')$  和  $(a', b')$  可用于下次迭代生产更优的解.

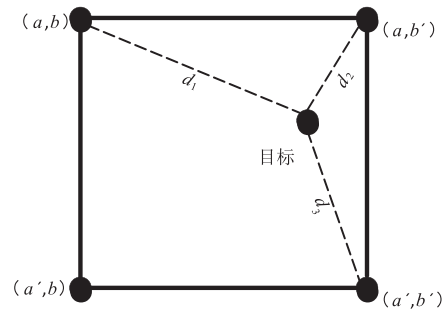


图1 反向学习策略在二维坐标系中的显示

### 3 仿真测试

#### 3.1 标准测试函数

为全面验证算法的优化性能,本文选择 Rashedi<sup>[12]</sup>等提出的已被广泛使用的 12 个标准函数进行寻优测试,测试函数如表 1 所示,其中  $f_1 \sim f_6$  是单峰函数,  $f_7 \sim f_9$  是维度可变的多峰函数,  $f_{10} \sim f_{12}$  是维度固定的多峰函数. 函数包括多个类型单峰函数适合测试算法的局部挖掘能力,多峰函数适合测试算法的全局的开拓能力,测试结果具有指导作用.

表 1 12 个测试函数简要信息

| 函数表达式  | 维数 | 搜索区间          | 最小值 |
|--|----|---------------|-----|
| $f_1(x) = \sum_{i=1}^n x_i^2$                                      | 30 | $[-100, 100]$ | 0   |
| $f_2(x) = \sum_{i=1}^n  x_i  + \prod_{i=1}^n  x_i $                | 30 | $[-10, 10]$   | 0   |
| $f_3(x) = \sum_{i=1}^n (\sum_{j=1}^i x_j)^2$                       | 30 | $[-100, 100]$ | 0   |
| $f_4(x) = \max_i \{ x_i , 1 \leq i \leq n\}$                       | 30 | $[-100, 100]$ | 0   |
| $f_5(x) = \sum_{i=1}^{n-1} [100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2]$ | 30 | $[-30, 30]$   | 0   |

续表

| 函数表达式  | 维数 | 搜索区间            | 最小值   |
|--|----|-----------------|-------|
| $f_6(x) = \sum_{i=1}^{n-1} ix_i^4 + \text{random}[0,1]$  | 30 | $[-1.28, 1.28]$ | 0     |
| $f_7(x) = \sum_{i=1}^{n-1} [x_i^2 - 10\cos(2\pi x_i) + 10]$  | 30 | $[-5.12, 5.12]$ | 0     |
| $f_8(x) = -20\exp\left(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^n x_i^2}\right) - \exp\left(\frac{1}{n}\sum_{i=1}^n \cos(2\pi x_i)\right) + 20 + e$                                     | 30 | $[-32, 32]$     | 0     |
| $f_9(x) = \frac{1}{4000}\sum_{i=1}^n x_i^2 - \prod_{i=1}^n \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$  | 30 | $[-600, 600]$   | 0     |
| $f_{10}(x) = \left(x_2 - \frac{5.1}{4\pi^2}x_1^2 + \frac{5}{\pi}x_1 - 6\right)^2 + 10\left(1 - \frac{1}{8\pi}\right)\cos x_1 + 10$   | 2  | $[-5, 5]$       | 0.398 |
| $f_{11}(x) = [1 + (x_1 + x_2 + 1)^2(19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2)] \times [30 + (2x_1 - 3x_2)^2 \times (18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2)]$ | 2  | $[-2, 2]$       | 3     |
| $f_{12}(x) = -\sum_{i=1}^4 c_i \exp\left(-\sum_{j=1}^3 a_{ij}(x_i - p_{ij})^2\right)$  | 3  | $[1, 3]$        | -3.86 |

### 3.2 实验分析

对于上述 12 个测试函数,为了保证公平,降低结果的偶然性,选择对比的 PSO、标准 GWO、TGWO<sup>[7]</sup>、OLGWO 四种算法的基本参数一致:初始种群规模为 30,最大迭代次数为 500,维度为 30 分别进行了 30 次独立测试,并将寻优结果的均值及标准差加以记录,实验结果如表 2 所示。

从表 2 中可看出,对于单峰函数的  $f_1 \sim f_4$ ,OLGWO 算法的精度最高,分别高出次优的 TGWO 算法 33 个、19 个、5 个和 9 个数量级,PSO 算法最差;对于单峰函数

$f_5$ ,OLGWO 算法虽准确度虽略低于标准 GWO,但标准差优于 GWO;对于单峰函数  $f_6$ ,OLGWO 算法精度略低于 TGWO,但明显优于标准 GWO. 对于维度可变的多峰函数  $f_7 \sim f_9$ ,OLGWO 算法精度的精度最优,且优于标准 GWO 约 2~3 个数量级. 对于固定维的多峰函数  $f_{10} \sim f_{12}$ ,OLGWO 算法的精度与标准 GWO 相当,但优于 TGWO. 综合三类函数分析,OLGWO 算法对于前两类函数精度提高的同时,对于第三类函数其精度仍较好,说明本文的方法能更好的平衡全局探索能力和局部精细搜索能力。

表 2 四种算法对测试函数的寻优结果

|       |     | PSO          | GWO          | TGWO          | OLGWO         |
|-------|-----|--------------|--------------|---------------|---------------|
| $f_1$ | 均值  | 1.822114e-04 | 9.734451e-28 | 2.030564e-197 | 5.562891e-230 |
|       | 标准差 | 2.506751e-04 | 1.585856e-27 | 0.000000e+00  | 0.000000e+00  |
| $f_2$ | 均值  | 3.066105e-02 | 9.507171e-17 | 2.049054e-99  | 1.167464e-118 |
|       | 标准差 | 4.284282e-02 | 9.289981e-17 | 4.454102e-100 | 1.259557e-118 |
| $f_3$ | 均值  | 8.056064e+01 | 1.106712e-05 | 2.278731e-196 | 8.247375e-201 |
|       | 标准差 | 2.908770e+01 | 2.059247e-05 | 0.000000e+00  | 0.000000e+00  |
| $f_4$ | 均值  | 1.033789e+00 | 1.085750e-06 | 1.068037e-98  | 6.745202e-107 |
|       | 标准差 | 2.425731e-01 | 1.442375e-06 | 1.683404e-99  | 7.530732e-107 |
| $f_5$ | 均值  | 6.594648e+01 | 2.703700e+01 | 2.898293e+01  | 2.791782e+01  |
|       | 标准差 | 3.915173e+01 | 9.038717e-01 | 1.546034e-02  | 3.272973e-01  |
| $f_6$ | 均值  | 1.803279e-01 | 2.069805e-03 | 8.345416e-05  | 1.099648e-04  |
|       | 标准差 | 6.574944e-02 | 8.292008e-04 | 8.265829e-05  | 8.441359e-05  |

续表

|          |     | PSO           | GWO           | TGWO          | OLGWO         |
|----------|-----|---------------|---------------|---------------|---------------|
| $f_7$    | 均值  | 5.770858e+01  | 2.787649e+00  | 1.845544e+02  | 0.000000e+00  |
|          | 标准差 | 1.231319e+01  | 4.483249e+00  | 1.151407e+02  | 0.000000e+00  |
| $f_8$    | 均值  | 1.432493e-01  | 1.035616e-13  | 1.924387e-14  | 4.440892e-15  |
|          | 标准差 | 3.867922e-01  | 1.802745e-14  | 6.123939e-15  | 0.000000e+00  |
| $f_9$    | 均值  | 9.133253e-03  | 2.736840e-03  | 0.000000e+00  | 0.000000e+00  |
|          | 标准差 | 1.028093e-02  | 6.541282e-03  | 0.000000e+00  | 0.000000e+00  |
| $f_{10}$ | 均值  | 3.978874e-01  | 3.979289e-01  | 1.018147e+00  | 3.998859e-01  |
|          | 标准差 | 0.000000e+00  | 1.632732e-04  | 1.010917e+00  | 2.521076e-03  |
| $f_{11}$ | 均值  | 3.000000e+00  | 3.000041e+00  | 3.620741e+01  | 3.000024e+00  |
|          | 标准差 | 1.569009e-15  | 4.509056e-05  | 3.794497e+01  | 3.740597e-05  |
| $f_{12}$ | 均值  | -3.862782e+00 | -3.862321e+00 | -3.424126e+00 | -3.857917e+00 |
|          | 标准差 | 2.597330e-15  | 1.072725e-03  | 3.625916e-01  | 2.807415e-03  |

表 2 中数据只给出了四种算法对 12 个测试函数寻找到的最优值,为了直观分析算法的收敛性,图 2 给出了四种算法对测试函数(由于篇幅所限,这里给出部分的)收敛对比图.图 2(a,b)是单峰函数  $f_2$  和  $f_6$  的收敛曲线,其中对于  $f_2$ ,OLGWO 收敛最快,对于  $f_6$ ,OLGWO 与 TGWO 收敛速度同为最快;图 2(c)~(e)是维度可变的多峰函数  $f_7 \sim f_9$  的收敛曲线,其中对于  $f_7$ ,TGWO 出现了明显下降甚至不收敛,OLGWO 仍保持了较好且很快的收敛速度,对于  $f_8$ ,OLGWO 表现最优,相对仍收敛最快,对于  $f_9$ ,OLGWO 的性能略次于 TGWO.图 2(f)~(h)是维度可变的多峰函数  $f_{10} \sim f_{12}$  的收敛曲线,三者的曲线基本相同,为了便于对照分析,  $f_{11}$  的收敛曲线图可认为展示了是  $f_{10}$  和  $f_{12}$  的迭代次数为 80 次时曲线的细节,可看出 OLGWO 与标准 GWO 保持了相对最优的收敛速度.由此收敛曲线的分析,可看出 OLGWO 算法收敛速度较快.

综合表 2 和图 2,OLGWO 兼顾了收敛速度、全局探索能力和局部精细搜索能力.

## 4 优化算法在交通流预测中的应用

### 4.1 OLGWO 算法优化 BP 网络模型的建立

传统的误差逆向传播(BP)神经网络在进行拟合时存在权值设置敏感等缺点,可通过调整神经元权值和阈值,利用优化后的 BP 神经网络进行训练和预测,可提高其预测精度,本文的详细步骤如下:

**步骤 1** 参数初始化.确定 BP 神经网络参数.根据实验数据确定 BP 神经网络的输入层节点数  $N_{input}$ 、隐含层节点数  $N_{hidden}$ 、输出层节点数  $N_{output}$  和激活函数;狼群数量  $N_{wolf}$ 、最大迭代次数  $Max_{iteration}$ 、个体狼维度  $dim$  及

其上限  $U_b$  和下限  $L_b$ .

**步骤 2** 确定适应度函数,将 BP 神经网络的观察值和预测值之差,公式表示为

$$f(x) = \sum_{i=1}^N |\hat{x}_i - x_i|$$

式中: $N$  为样本数,  $x_i$  和  $\hat{x}_i$  分别为观察值及预测值.

**步骤 3** 根据  $N_{wolf}$ 、 $dim$ 、 $U_b$  和  $L_b$  初始化狼种群.

**步骤 4** 根据步骤 2 的公式计算并记录种群个体的适应度,根据适应度选择出  $\alpha$ 、 $\beta$ 、 $\delta$  和  $\omega$  狼.

**步骤 5** 以  $\alpha$  狼作为目标,通过计算每个  $\omega$  狼与目标的 Spearman 相关系数发现  $\omega$  狼中方向相反的  $\omega_{contrary}$  狼( $r_s < 0$ ),对  $\omega_{contrary}$  狼应用反向学习策略产生反向  $\omega_{opposite}$  狼,采用这些狼对 BP 网络进行训练与预测,按照步骤 2 的公式计算适应度函数,并将适应度小的狼用于更新狼种群.

**步骤 6** 更新  $a$ ,并按照式(3)和(4)随机生成  $C$  和  $A$ .

**步骤 7** 按照式(5)~(11)完成种群的更新.

**步骤 8** 如果没有达到最大迭代次数  $Max_{iteration}$ ,重复至步骤 4,达到则保存记录最优解继续下一步.

**步骤 9** 将最优解赋值给 BP 网络,并随机生成训练数据和测试数据,记录每次的输出结果.

### 4.2 评价指标

本文选取平均绝对百分比误差(Mean Absolute Percentage Error, MAPE)和均方根误差(Root Mean Square Error, RMSE)作为评价指标.其中 MAPE 侧重评价算法对实验数据的预测精度, RMSE 侧重评价算法对实验数据的预测鲁棒性.

$$MAPE = \frac{1}{N} \sum_{i=1}^N \frac{|x_i - \hat{x}_i|}{x_i} \quad (14)$$

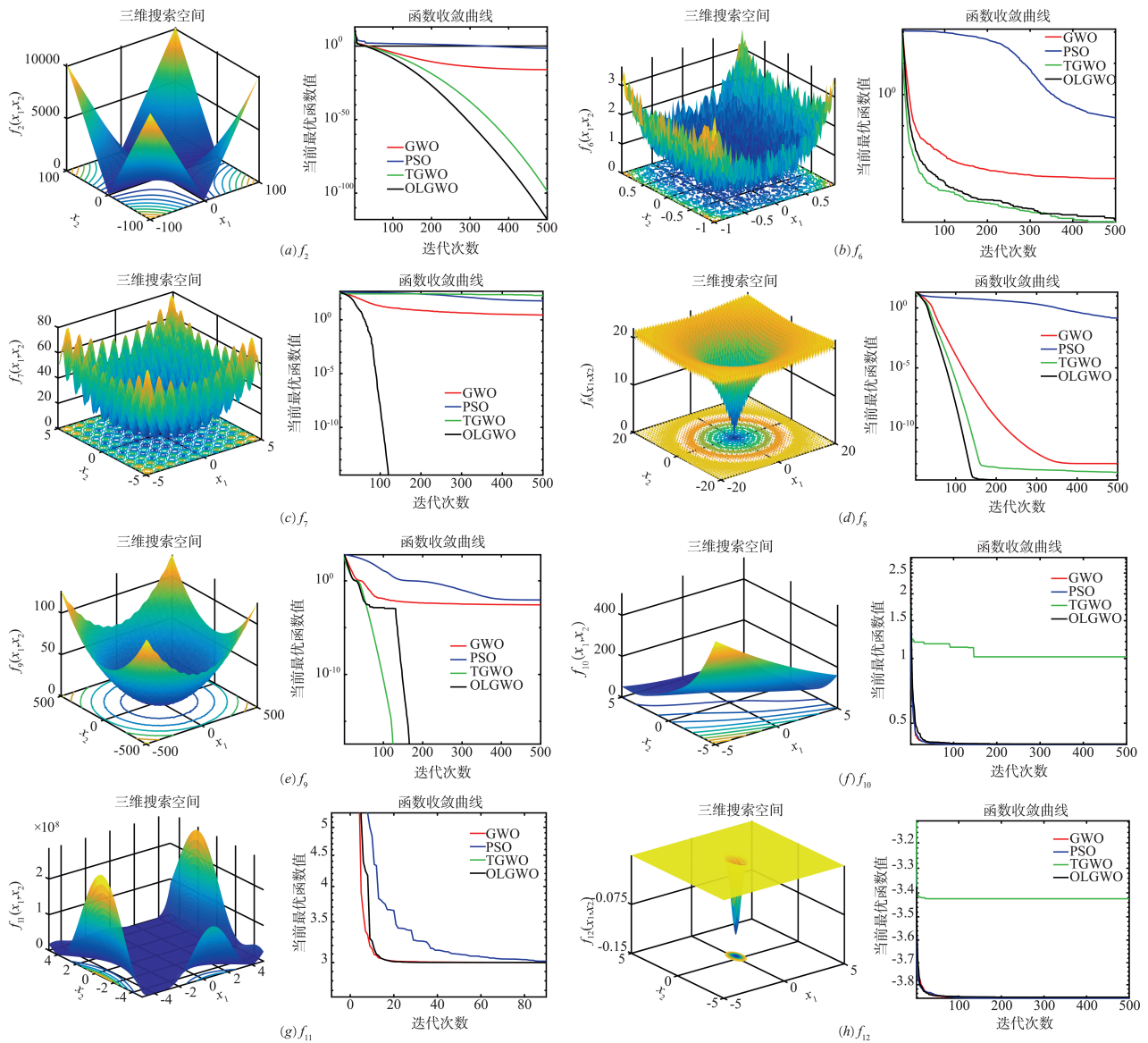


图2 PSO, GWO, TGWO和OLGWO算法对多种测试函数的收敛曲线

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N |x_i - \hat{x}_i|^2} \quad (15)$$

式中,  $x_i$  和  $\hat{x}_i$  观察值及预测值.

### 4.3 预测结果对比分析

本文采用的数据是美国加州 2014 年 5 月 1 日 ~ 6 月 30 日 24 条路段每 5min 的工作日流量数据,原始数据中为零数据已取其前后数据均值进行了修复,数据中传感器编号即视为路段编号. 实验中假设编号为 760080 的传感器流量为标签数据,编号为 767838、773656、760074、71855 和 760226 五个传感器流量为属性数据,连续取 1000 组数据,其中 800 组用于训练,200 组用于测试. 为了公平比较各算法的性能,各模型的参数保持基本一致. BP 网络的输入层、隐含层、输出层为

节点数均为 [5 6 1], 隐藏层和输出层的激活函数分别为 tansig 函数和 purelin 函数. 四种优化方法 (PSO-BP、GWO-BP、TGWO-BP、OLGWO-BP) 每次独立运行 50 次,取平均值作为最终预测结果. 本文给出在种群数量为 10 时,不同迭代次数下的预测结果,如表 3 所示.

对于无缺失数据,迭代 50 次时,OLGWO 的 MAPE 仅比 PSO 高 0.02%,且 OLGWO 的 RMSE 最小;迭代 100 次时,OLGWO 的 MAPE 和 RMSE 均为最小,TGWO 为次优;迭代 200 次时,OLGWO 比 GWO 的 MAPE 略大,但 RMSE 最小;表明 OLGWO 算法在收敛速度、寻优精度方面相对 GWO 和 TGWO 有改进,同时鲁棒性表现更好.

为了进一步验证改进效果,本文给出了不同缺失率下的 MAPE 和 RMSE. 缺失率 10% 迭代 200 次时,OL-

GWO 比最优的 GWO 和次优的 TGWO 的 MAPE 仅小 0.31%,但 OLGWO 的 RMSE 最小,表明 OLGWO 的寻优能力与 GWO 和 TGWO 相当,但鲁棒性表现更好;缺失率 20% 迭代 200 次时 OLGWO 的 MAPE 比 GWO 和 TGWO 分别高 0.5% 和 0.86%,迭代 300 次时,OLGWO

的 MAPE 分别比 GWO 和 TGWO 分别高出 1.95% 和 1.71%,再次表明 OLGWO 在寻优精度上比 GWO 和 TGWO 更好;缺失率 30% 数据质量较差,在迭代 300 次时,OLGWO 的 MAPE 与 GWO 相当,但 OLGWO 的 RMSE 最小,表明 OLGWO 的稳定性更好。

表 3 PSO、GWO、TGWO、OLGWO 优化 BP 网络的预测结果

| 缺失率 | 迭代次数  | 评价指标 | PSO-BP       | GWO-BP       | TGWO-BP | OLGWO-BP      |
|-----|-------|------|--------------|--------------|---------|---------------|
| 无   | 50 次  | MAPE | <b>5.61%</b> | 5.74%        | 5.70%   | 5.63%         |
|     |       | RMSE | 22.497       | 22.938       | 23.076  | <b>21.916</b> |
| 无   | 100 次 | MAPE | 6.01%        | 5.72%        | 5.66%   | <b>5.59%</b>  |
|     |       | RMSE | 23.096       | 23.443       | 22.691  | <b>21.900</b> |
| 无   | 200 次 | MAPE | 7.69%        | <b>5.69%</b> | 6.09%   | 5.76%         |
|     |       | RMSE | 23.468       | 22.336       | 22.536  | <b>22.156</b> |
| 10% | 200 次 | MAPE | 18.19%       | <b>6.81%</b> | 7.07%   | 7.12%         |
|     |       | RMSE | 44.059       | 25.135       | 27.518  | <b>24.598</b> |
| 20% | 200 次 | MAPE | 12.92%       | 8.24%        | 8.60%   | <b>7.74%</b>  |
|     |       | RMSE | 44.939       | 33.131       | 28.850  | <b>28.296</b> |
| 20% | 300 次 | MAPE | 12.95%       | 10.00%       | 9.76%   | <b>8.05%</b>  |
|     |       | RMSE | 43.879       | 38.945       | 31.904  | <b>29.557</b> |
| 30% | 300 次 | MAPE | 14.12%       | 12.84%       | 16.04%  | <b>12.06%</b> |
|     |       | RMSE | 44.499       | 44.728       | 50.067  | <b>37.950</b> |

适应度函数的定义对于智能启发算法应用至关重要,直接决定收敛结果<sup>[13]</sup>。算法性能一般随着迭代次数的增多精度会有不同程度的提高,但计算时间也随之增加。对于工程应用,可通过适应度函数收敛曲线确定合适的迭代次数。本文的应用属回归问题,将绝对误差作为适应度函数,四种方法(PSO、GWO、TGWO、OLGWO)优化 BP 网络(50 次取平均值)的适应度函数的收敛曲线如图 3 所示,横坐标为迭代次数,纵坐标为适应度函数值,不同的应用误差大小不一,为更具通用性,图 3 中纵坐标未标注具体值。图 3 中整体看,GWO 和 PSO 较差,TGWO 和 OLGWO 较好。特别对于 OLGWO,迭代次数为 130、150、180 次是仍有不同程度降低。分析图 3 可得,迭代次数可以选择在 130~180 之间。对于分

类问题的应用,可以根据实际需求构造准确率、错误率、召回率、F 值等适应度函数用于优化模型的构建<sup>[14,15]</sup>。

### 5 结论

本文在经典灰狼算法迭代过程中,采用 Spearman 相关系数结合反向学习策略更新种群,提高了算法的全局探索能力和局部精细搜索能力。通过对 12 个基准测试函数的寻优实验,结果表明 OLGWO 较经典的 GWO、PSO 以及新近的 TGWO 有更好的全局探索和局部精细搜索的平衡能力,同时有较好的收敛速度。采用美国加州交通流数据,通过四种方法(OLGWO、TGWO、GWO、PSO)优化 BP 网络模型预测交通流,经 OLGWO 算法优化后的预测结果比另三种算法最高分别有 1.95%、3.98% 和 11.07% 的提升,在交通流工程领域取得了满意的应用效果。神经网络的网络层数、每层的节点数、层的激活函数等如果均进行优化,仍是一项非常耗时的工作,寻求收敛速度更快的策略是下一步的研究方向。

### 参考文献

[1] 代亮,梅洋,钱超,等.基于深度学习的短时交通量预测研究综述[J].计算机科学,2019,46(3):39-47.

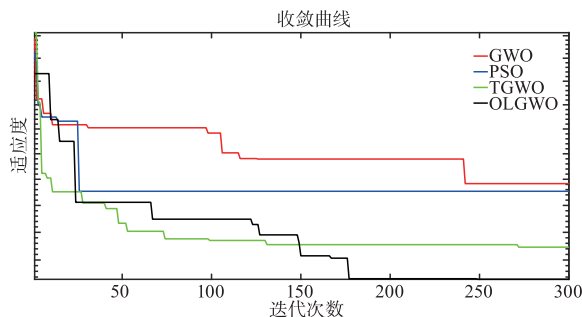


图 3 PSO、GWO、TGWO、OLGWO 优化 BP 网络的适应度函数收敛曲线

- Dai L, Mei Y, Qian C, et al. Survey on short-term traffic flow forecasting based on deep learning[J]. Computer Science, 2019, 46(3):39-47. (in Chinese)
- [2] Moretti F, Pizzuti S, Panzieri S, et al. Urban traffic flow forecasting through statistical and neural network bagging ensemble hybrid modeling[J]. Neurocomputing, 2015, 167(C):3-7.
- [3] Zhao Z, Chen W H, Wu X M, et al. LSTM network: a deep learning approach for short-term traffic forecast[J]. IET Intelligent Transport Systems, 2017, 11(2):68-75.
- [4] 蔡亮, 周泓岑, 白恒, 等. 基于多层 BiLSTM 和改进粒子群算法的应用负载预测方法[J]. 浙江大学学报(工学版), 2020, 54(12):2414-2422.
- Cai L, Zhou H C, Bai H, et al. Application load forecasting method based on multi-layer bidirectional LSTM and improved PSO algorithm[J]. Journal of Zhejiang University (Engineering Science), 2020, 54(12):2414-2422. (in Chinese)
- [5] Mirjalili S, Mirjalili S M, Lewis A. Grey wolf optimizer[J]. Advances in Engineering Software, 2014, 69(3):46-61.
- [6] 傅蔚阳, 刘以安, 薛松. 基于灰狼算法与小波神经网络的目标威胁评估[J]. 浙江大学学报(工学版), 2018, 52(4):680-686.
- FU W Y, LIU Y A, XUE S. Target threat assessment using grey wolf optimization and wavelet neural network[J]. Journal of Zhejiang University (Engineering Science), 2018, 52(4):680-686. (in Chinese)
- [7] 张文胜, 郝孜奇, 朱冀军, 等. 基于改进灰狼算法优化 BP 神经网络的短时交通流预测模型[J]. 交通运输系统工程与信息, 2020, 20(2):196-203.
- Zhang W S, Hao Z Q, Zhu J J, et al. BP neural network model for short-time traffic flow forecasting based on transformed grey wolf optimizer algorithm[J]. Journal of Transportation Systems Engineering and Information Technology, 2020, 20(2):196-203. (in Chinese)
- [8] Mee Song H, Sulaiman M H, Mohamed M R. An application of grey wolf optimizer for solving combined economic emission dispatch problems[J]. International Review on Modelling and Simulations (IREMOS), 2014, 7(5):838.
- [9] Heidari A A, Pahlavani P. An efficient modified grey wolf optimizer with Lévy flight for optimization tasks[J]. Applied Soft Computing, 2017, (60):115-134.
- [10] Tizhoosh H R. Opposition-based learning: a new scheme for machine intelligence[A]. International Conference on Computational Intelligence for Modelling, Control and Automation, and International Conference on Intelligent Agents, Web Technologies and Internet Commerce[C]. Vienna, Austria: IEEE Computer Society, 2005. 695-701.
- [11] 樊嵘, 孟大志, 徐大舜. 统计相关性分析方法研究进展[J]. 数学建模及其应用, 2014, 3(1):1-12.
- Fan R, Meng D Z, Xu D S. Survey of research process on statistical correlation analysis[J]. Mathematical Modeling and Its Applications, 2014, 3(1):1-12. (in Chinese)
- [12] Rashedi E, Nezamabadi-pour, H Saryazdi S. GSA: A gravitational search algorithm[J]. Information Sciences, 2009, 179(13):2232-2248.
- [13] 沈焱萍, 郑康锋, 伍淳华, 等. 智能启发算法在机器学习中的应用研究综述[J]. 通信学报, 2019, 40(12):124-137.
- Shen Y P, Zheng K F, Wu C H, et al. Survey of research on application of heuristic algorithm in machine learning[J]. Journal on Communications, 2019, 40(12):124-137. (in Chinese)
- [14] Ali M H, Al Mohammed B A D, Ismail A, et al. A new intrusion detection system based on fast learning network and particle swarm optimization[J]. IEEE Access, 2018, 6:20255-20261.
- [15] Wu Q, He F F, Fan X M. The intelligent control system of traffic light based on fog computing[J]. Chinese Journal of Electronics, 2018, 27(6):1265-1270.

## 作者简介



张兴辉 男, 1982 年 4 月生于辽宁省营口市. 2009 年获长春理工大学硕士学位, 其后在安康学院从事教学工作, 讲师. 2018 年至今在西安理工大学信息与控制系统攻读博士学位. 主要研究方向为车联网、智能交通控制.

E-mail: 1180311021@stu.xaut.edu.cn



樊秀梅(通讯作者) 女, 1967 年 7 月生于山西省山阴县. 博士, 西安理工大学自动化与信息工程学院教授, 博士生导师. 主要研究领域包括车联网、下一代移动互联网、自动驾驶网联协同控制、边缘计算等.

E-mail: xmfan@xaut.edu.cn