

文章编号:2095-6134(2020)05-0673-08

基于改进粒子群的 BP 神经网络 WSN 数据融合算法^{*}

王虹^{1,2†}, 徐佑宇^{1,2}, 谭冲¹, 刘洪¹, 郑敏¹

(1 中国科学院上海微系统与信息技术研究所, 上海 200050; 2 中国科学院大学, 北京 100049)

(2019 年 3 月 5 日收稿; 2019 年 5 月 8 日收修改稿)

Wang H, Xu Y Y, Tan C, et al. Information fusion algorithm based on improved particle swarm BP neural network in WSN[J]. Journal of University of Chinese Academy of Sciences, 2020,37(5):673-680.

摘 要 针对无线传感器网络(WSN)数据融合算法中传统反向传播(BP)神经网络收敛速度慢、对初值敏感和易陷入局部最优解的问题,提出基于改进粒子群的 BP 神经网络 WSN 数据融合算法(BSO-BP)。用天牛须搜索(BAS)算法对粒子群算法进行改进,利用改进后的粒子群算法优化 BP 神经网络权值和阈值,引入 WSN 数据融合中,簇首节点通过优化训练后的 BP 神经网络对采集数据进行特征提取,将融合后的数据发送至汇聚节点。仿真实验表明,BSO-BP 算法能有效地提高融合精度和收敛速度,减少冗余数据传输,延长网络生命周期。相较于传统 BP 数据融合算法和 PSO-BP 算法,BSO-BP 算法减少了至少 11% 的平均相对误差和 13.89% 的均方根误差。

关键词 无线传感器网络;数据融合;BP 神经网络;粒子群算法;天牛须搜索算法

中图分类号: TN929.5 **文献标志码:** A **doi:** 10.7523/j.issn.2095-6134.2020.05.012

Information fusion algorithm based on improved particle swarm BP neural network in WSN

WANG Hong^{1,2}, XU Youyu^{1,2}, TAN Chong¹, LIU Hong¹, ZHENG Min¹

(1 Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, Shanghai 200050, China; 2 University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract Back-propagation (BP) neural network has low convergence speed, is sensitive to the initial value, and easily falls into the local optimal solution in data fusion algorithms in wireless sensor network (WSN). To solve these problems, a data fusion algorithm based on improved particle swarm optimization BP neural network in WSN (BSO-BP) is proposed. The beetle antennae search (BAS) algorithm is used to improve the particle swarm optimization. Then the improved particle swarm optimization is used to optimize the BP neural network weights and thresholds, which are applied to WSN data fusion. The cluster head nodes extract the feature of the collected data by optimizing the trained BP neural network, and send the merged data to the sink node. Simulation

^{*} 国家自然科学基金(61401445)资助

[†] 通信作者, E-mail: hong.wang@mail.sim.ac.cn

results show that BSO-BP algorithm effectively improves the fusion accuracy and convergence speed, decreases the redundant data communication and prolongs network lifetime. BSO-BP algorithm reduces the relative error by 12.4% and the root-mean-square error by 11%, compared to BP and PSO-BP algorithms.

Keywords wireless sensor networks (WSN); data fusion; BP neural network; particle swarm optimization; beetle antennae search algorithm

无线传感器网络 (wireless sensor network, WSN) 是由大量小尺寸、低成本的微传感器节点通过自组织和多跳方式形成的通信网络^[1], 将传感器数据发送给汇聚节点以实现对特定区域的环境信息的探测和感知。WSN 中传感器节点能量有限, 并且节点电池由于环境因素难以替换^[2]。WSN 中传感器节点在监测区域分布较密, 会导致监测区域的重叠, 产生的大量冗余数据传输至汇聚节点会造成额外的能量消耗, 缩短整个网络的生命周期^[3]。对 WSN 数据进行数据融合处理, 能有效地减少冗余数据, 降低传感器节点能耗, 延长网络生命周期^[4]。

针对无线传感网数据传输能耗较高的问题, 科研人员提出各种 WSN 数据融合算法来减少数据传输, 以降低传感器节点能耗^[5]。文献[3]设计一种基于反向传播 (back propagation, BP) 神经网络的数据融合算法 (back-propagation networks data aggregation, BPNDA), 该算法将 BP 神经网络和传感器网络分簇路由协议有机结合, 将每个簇设计成一个神经网络模型, 通过神经网络提取对原始数据进行特征提取, 然后将特征数据发送给汇聚节点, 从而提高数据收集效率, 延长网络生存时间。但是 BP 神经网络对于初始权值和阈值很敏感, 易陷入局部极小值。文献[6]利用遗传算法对 BP 神经网络的权值和阈值参数进行优化, 提出一种基于遗传算法 (genetic algorithm, GA) 的 BP 神经网络数据融合算法 GABP 算法, 能有效减少冗余数据, 延长网络生命周期。但遗传算法存在搜索能力较差, 求解速度较慢的缺陷。文献[7]把蚁群算法应用到 BP 神经网络, 对其结构及参数进行优化, 以克服 BP 神经网络容易陷入局部最优解的缺点。利用优化后的 BP 神经网络进行数据融合, 可以有效提取特征数据, 延长网络生命周期。文献[8]提出一种基于粒子群 (particle swarm optimization, PSO) 优化 BP 神经网络的无线传感器网络数据融合算法 PSO-BP, 利用粒子群算法优化 BP 神经网络参数, 将优化后

的 BP 神经网络和传感器网络分簇路由协议有机结合, 有效提升数据融合效率, 延长网络生命周期。但是由于粒子群算法种群搜索空间中多样性的丢失, 所以粒子群算法存在陷入局部最优解的缺点^[9]。

综上所述, 传统 BP 神经网络应用于无线传感器网络中存在训练困难、易陷入局部最优解的缺陷; GABP 算法寻优效果受种群初始分布影响较大, 求解速度缓慢; PSO-BP 算法中粒子群算法存在早熟收敛、维数灾难、缺乏多样性、易于陷入局部极值等问题, 仍具有较大优化空间。本文以优化寻优性能和融合精度为目标, 提出一种基于天牛须搜索 (beetle antennae search, BAS) 算子改进粒子群算法的 WSN 数据融合算法 BSO-BP 算法, 通过天牛须搜索算子优化粒子群算法^[10], 增加粒子群算法的搜索空间多样性, 丰富粒子群算法的寻优空间以避免陷入局部最优解。利用优化后的粒子群算法对 BP 神经网络的初始权值和阈值参数进行优化, 得到的全局较优解作为 BP 神经网络结构参数的初始值并进行进一步优化训练, 从而克服 BP 神经网络训练极易陷入局部最优解的缺陷, 避免传统 BP 神经网络初始权值和阈值参数随机选取导致的网络震荡, 提高 BP 神经网络的泛化能力和收敛能力。将 BSOBP 算法运用到无线传感器网络分簇路由数据融合中, 簇首节点利用优化后的 BP 神经网络对传感器采集数据进行数据融合提取特征数据, 数据精度具有较高提升, 将融合后的数据发送至基站节点, 有效减少冗余数据传输, 延长网络生命周期。

1 改进 PSO 算法

1.1 粒子群算法

粒子群 (PSO) 算法是一种基于鸟类群体的社会行为的群体优化算法, 通过群体信息共享机制寻找当前最优值以获取全局最优值。每个粒子的行为模仿鸟类飞行, 粒子位置根据自身的惯性速

度、历史经验以及群体社会经验来进行寻优更新^[11],寻优公式如下

$$\mathbf{v}_{id}^{k+1} = w\mathbf{v}_{id}^k + c_1r_1(\mathbf{pbest}_{id}^k - \mathbf{x}_{id}^k) + c_2r_2(\mathbf{gbest}_{id}^k - \mathbf{x}_{id}^k), \quad (1)$$

$$\mathbf{x}_{id}^{k+1} = \mathbf{x}_{id}^k + \mathbf{v}_{id}^{k+1}. \quad (2)$$

式中: $d = 1, 2, \dots, D$ (D 表示粒子总维度), $i = 1, 2, \dots, N$ (N 为粒子群粒子数量), k 为当前迭代次数, \mathbf{v}_{id}^k 表示第 k 次迭代时粒子 i 的第 d 维位置矢量, \mathbf{x}_{id}^k 表示第 k 次迭代时粒子 i 的第 d 维位置矢量, w 为惯性权重, c_1 和 c_2 为加速度常数, r_1 和 r_2 为两个取值范围为 $[0, 1]$ 的随机函数。 \mathbf{pbest}_{id}^k 表示粒子 i 的历史最优适应度在第 k 次迭代中的第 d 维位置, \mathbf{gbest}_{id}^k 表示群体粒子的历史最优适应度在第 k 次迭代中的第 d 维位置矢量。

1.2 基于天牛须搜索算法对粒子群的改进

天牛须搜索 (BAS) 算法是由 Jiang 和 Li 在 2018 年提出的一种新型启发式算法^[12], 它是一项基于天牛觅食原理的适用于目标函数优化的新技术, 其灵感来源于天牛觅食过程: 天牛觅食时, 通过食物气味强弱来觅食。天牛有两只灵敏的长触角, 如果左触角收到的气味强度大于右边, 那下一步天牛就会往左边飞, 反之则往右边飞。依据这一简单原理天牛就可以有效找到食物, BAS 对函数具体形式没有要求, 不需要梯度信息, 就可以实现高效寻优。BAS 收敛快, 并且算法简单容易实现, 容易移植到单片机运行, 很适合实时性要求较高的场景。天牛须搜索算法模型主要包括搜索行为和检测行为:

1) 创建天牛须朝向的归一化随机向量

$$\mathbf{b} = \frac{\text{rands}(m, 1)}{\|\text{rands}(m, 1)\|}, \quad (3)$$

式中: $\text{rands}()$ 为随机函数; m 表示空间维度。

2) 创建天牛左右须空间搜索模型, 模拟天牛须行为

$$\begin{cases} x_{rk} = x^k + d^k \times \mathbf{b}/2 \\ x_{lk} = x^k - d^k \times \mathbf{b}/2 \end{cases}, \quad (4)$$

式中: x_{rk} 表示天牛右须在第 k 次迭代时的位置坐标; x_{lk} 表示天牛左须在第 k 次迭代时的位置坐标; x^k 表示天牛在第 k 次迭代时的质心坐标; d^k 表示左右两须之间的搜索距离。

3) 根据适应度函数判断左右须气味强度差异, 即计算 $f(x_{lk})$ 和 $f(x_{rk})$ 的强度, f 为适应度

函数。

4) 迭代更新天牛的位置

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \delta^k \times \mathbf{b} \times \text{sign}(f(x_{rk}) - f(x_{lk})), \quad (5)$$

式中: δ^k 表示在第 k 次迭代时的步长因子; $\text{sign}()$ 为符号函数。

5) 搜索距离和步长更新

$$\begin{aligned} d^k &= \eta_d d^{k-1} + d_0, \\ \delta^k &= \eta_\delta \delta^{k-1} + \delta_0. \end{aligned} \quad (6)$$

式中: η_d 代表搜索距离的衰减更新系数; η_δ 代表步长的衰减更新系数; d_0 和 δ_0 分别是最小距离阈值和最小步长阈值, 避免 d^k 和 δ^k 为 0。

天牛须搜索算法的搜索行为对于 PSO 算法有很大作用, 赋予 PSO 算法中粒子天牛搜索行为特性, 可以有效增加粒子种群变化的多样性, 克服 PSO 算法寻优过程中易陷入局部最优的问题。本文根据基于适应度值的轮盘赌算法选取多个粒子进行天牛须搜索, 每次迭代选取赋予天牛搜索行为的粒子个数为

$$N_i = \text{swarmize} - \left(\frac{\text{iter}}{\text{maxgen}} \right) (\text{swarmize} - 2), \quad (7)$$

式中: swarmize 为粒子种群粒子总数, iter 和 maxgen 分别为当前迭代次数和总迭代次数。粒子群搜索初期, 具有天牛搜索行为的粒子能使种群更具多样性, 扩展了寻找最优解的范围, 避免陷入局部最优解, 随着迭代次数增加, 粒子群越来越收敛, 此时应减少具有天牛搜索行为的粒子数量, 有利于减少计算量, 快速收敛。具有天牛搜索行为的粒子根据式 (3) 随机生成天牛须朝向向量, 并对式 (4) 进行修改确定天牛须左右须坐标

$$\begin{cases} \mathbf{x}_{ird}^k = \mathbf{x}_{id}^k + \mathbf{v}_{id}^k \times d^k/2 \\ \mathbf{x}_{ild}^k = \mathbf{x}_{id}^k - \mathbf{v}_{id}^k \times d^k/2 \end{cases}, \quad (8)$$

式中: \mathbf{x}_{ird}^k 和 \mathbf{x}_{ild}^k 分别表示第 k 次迭代时具有天牛搜索行为的粒子 i 的左右须第 d 维位置矢量。

引入粒子天牛须搜索行为增量

$$\xi_{id}^{k+1} = \delta^k \times \mathbf{v}_{id}^k \times \text{sign}(f(x_{ird}^k) - f(x_{ild}^k)), \quad (9)$$

式中: ξ_{id}^{k+1} 代表第 $k+1$ 次迭代时的搜索行为增量。对粒子群的寻优计算式进行改进, 添加天牛须搜索行为增量

$$\mathbf{v}_{id}^{k+1} = w\mathbf{v}_{id}^k + c_1r_1(\mathbf{pbest}_{id}^k - \mathbf{x}_{id}^k) + c_2r_2(\mathbf{gbest}_{id}^k - \mathbf{x}_{id}^k), \quad (10)$$

$$\mathbf{x}_{id}^{k+1} = \mathbf{x}_{id}^k + (1 - \lambda)\mathbf{v}_{id}^{k+1} + \lambda\xi_{id}^{k+1}. \quad (11)$$

式中: w 为惯性权重, 反映粒子的运动习惯, 理论上, w 取值较小时局部搜索能力较强, w 取值较大可以增强全局搜索能力。文献[13]分析比较多种动态调整惯性权重的方法, 发现对于多峰函数优化, 凸函数递减法综合来说收敛速度最快, 求解精度较高。因此本文惯性权重 w 采用凸函数递减法

$$w = (w_{\max} - w_{\min}) \left(1 - \frac{\text{iter}}{\text{maxgen}}\right)^3 + w_{\min}. \quad (12)$$

式中: w_{\max} 和 w_{\min} 分别为惯性权重上下限值, iter 和 maxgen 分别为当前迭代次数和总迭代次数。 c_1 和 c_2 为学习因子, c_1 表示粒子对自身历史经验的学习, c_2 表示粒子对共享的群体历史经验的学习^[14], PSO 算法运行初期, 侧重于粒子个体的探索认知, 末期需要增强对群体经验的学习能力, 因此, c_1 和 c_2 分别随着迭代次数递减和递增:

$$c_1 = c_m + \frac{\text{iter}}{\text{maxgen}}(c_n - c_m), \quad (13)$$

$$c_2 = c_n + \frac{\text{iter}}{\text{maxgen}}(c_m - c_n). \quad (14)$$

式中: c_m 和 c_n 为常数, 满足 $c_m > c_n$ 。

改进 PSO 算法中, λ 为天牛搜索行为增量的权重因子, 搜索前期, 天牛搜索行为增量权重较大有利于增加全局搜索能力, 后期随着具有天牛搜索行为的粒子数量的减少, 天牛搜索行为增量权重也应当降低, 因此搜索行为增量权重因子 λ 的赋值如下

$$\lambda = (\lambda_{\max} - \lambda_{\min}) \left(1 - \frac{\text{iter}}{\text{maxgen}}\right) + \lambda_{\min}. \quad (15)$$

式中: λ_{\max} 和 λ_{\min} 为常数, 分别代表增量权重因子的上下限值。

2 BSO-BP 数据融合算法

2.1 算法网络模型

WSN 多传感器数据融合算法可分为数据级融合、特征性融合和决策级融合 3 类。决策级融合通过不同种类的传感器观测同一个目标, 例如温度、湿度、风速等传感器感知森林某区域环境信息, 每个传感器在本地完成包括预处理、特征抽取、识别或判决等基本处理, 以建立对所观察目标的初步结论。然后通过关联处理进行决策层融合判决, 最终获得联合推断结果。决策级融合包括贝叶斯估计法、专家系统、神经网络法、模糊集理

论等, 本文对决策级融合中的神经网络算法进行研究改进。

神经网络是由大量具有运算能力的单元广泛互联而组成的, 通过高效的学习机制模拟人脑信息处理的过程。神经网络由于其强大的计算能力, 因此在信号处理、计算机视觉、自然语言优化处理等领域应用广泛^[15]。BP 神经网络是一种典型的多层前向型神经网络^[16]。BP 神经网络由输入层、隐藏层、输出层组成, 它利用学习机制来存储输入和输出的映射关系, 其权值和阈值参数通常采用反向传播的策略, 借助最速梯度信息来获得使网络误差最小化的参数组合, 有很强的非线性映射能力、自学习和自适应能力、泛化能力和容错能力。将 BP 神经网络模型利用到 WSN 中处理数据融合问题, 能有效降低数据特征维数, 提高数据融合效率和精度, 从而减少节点能耗。利用 BP 神经网络结合无线传感器分簇路由算法构建 WSN 数据融合模型, 网络结构如图 1 所示。

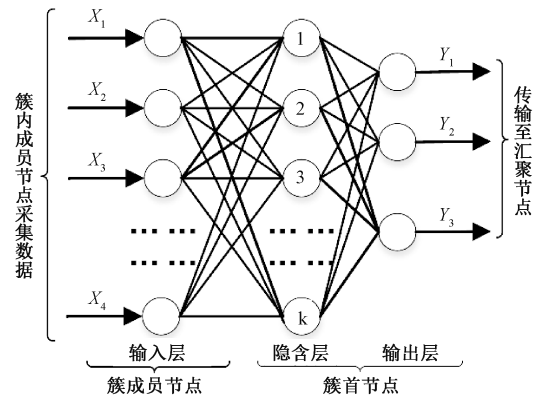


图 1 BSO-BP 算法神经网络结构

Fig. 1 Neural network architecture of BSO-BP algorithm

按照 LEACH (low energy adaptive clustering hierarchy) 分簇路由算法完成分簇后, 簇内传感器节点首先对监测到的原始数据进行滤波等预处理, 然后将数据传输至簇首节点, 簇首节点利用 BP 神经网络对数据进行融合, 提取特征数据, 减少冗余数据, 将融合后的特征数据传输至汇聚节点, 从而减少数据传输量, 延长网络生命周期^[17]。但传统 BP 神经网络也存在一定缺陷, 初始权值参数的选择具有随机性, 网络对初始连接权值的选择很敏感, 初值选取不当会导致训练结果不佳, 易陷入局部最优解。

2.2 算法步骤

利用改进后的粒子群算法对 BP 神经网络数据融合算法进行优化, 并应用于 LEACH 分簇路

由算法中,提出一种新型智能数据融合算法——BSO-BP 算法。

LEACH 分簇路由算法中需要定期更新簇首以平衡能耗,相应地簇集也需动态更新,簇首的变化会导致簇内 BP 神经网络结构的变化,BP 神经网络进行数据融合需要权值和阈值参数,因此簇首更新后,在 WSN 进入自治工作状态前,BP 神经网络的参数也需更新。WSN 中传感器节点运算能力、节点能量以及存储能力有限,因此参数的训练更新在基站节点进行,BSO-BP 算法整体步骤如下:

步骤 1: 每轮簇首节点更新,分簇形成后,簇首记录簇内节点信息并传输至基站节点。

步骤 2: 基站根据簇集信息确定 BP 神经网络结构。

步骤 3: 基站选取与该分簇相匹配的样本数据集,结合改进后的 PSO 算法对 BP 神经网络进行训练优化,以得到对应簇集的 BP 神经网络连接权值和阈值参数。

步骤 4: 基站将相应簇集的 BP 神经网络连接权值和阈值参数发送至各簇集节点内,存储参数供下次使用。

步骤 5: 每个簇集根据接收到的参数报文构建对应的 BP 神经网络,簇首节点对簇内成员节点上传的数据进行融合处理并将结果传输至基站节点。至此,完成一轮分簇路由通信过程。

上述步骤 3 中改进 PSO 算法优化 BP 神经网络算法的目的是优化 BP 神经网络的权值阈值参数。利用改进 PSO 对 BP 神经网络的连接权值和阈值参数进行优化,获得一个全局优化解。将全局优化解作为 BP 神经网络的初始权值和阈值参数,BP 神经网络进一步训练优化,以获得全局最优参数。这样,克服了 BP 神经网络由于初始权值选取的随机性导致的收敛性不佳、训练时间较长、易陷入局部最优解等缺陷,提高了 BP 神经网络的收敛能力和求解精度,改进 PSO 算法优化 BP 神经网络算法具体实现步骤如下:

Step 1: 确定搜索空间维度。WSN 分簇形成后,每个簇内成员节点个数即是输入层神经元个数 M ,隐藏层神经元个数为 N ,输出层神经元个数为 1。因此,搜索空间维度 $D = M \times N + N \times 1 + N + 1$ 。

Step 2: 初始化设置。初始化 n 个粒子的初始位置向量 $\mathbf{x}_i (i = 1, 2, \dots, n)$, 其中 $\mathbf{x}_i = (x_{i1},$

$x_{i2}, \dots, x_{iD})$ 为第 i 个粒子的 D 维搜索向量,包含 BP 神经网络中输入层和隐藏层的权值,隐藏层和输出层的权值,以及隐藏层和输出层的阈值;初始化 n 个粒子的速度向量 $\mathbf{v}_i (i = 1, 2, \dots, n)$, 其中 $\mathbf{v}_i = (v_{i1}, v_{i2}, \dots, v_{iD})$ 为第 i 个粒子的速度向量;初始化速度变化范围 $[-V_{\max}, V_{\max}]$;初始化最大迭代次数 K 。

Step 3: 定义适应度函数。对样本数据进行融合,根据样本数据的均方根误差 (RMSE) 作为适应度函数

$$f(\mathbf{x}_i) = \sqrt{\frac{1}{m} \sum_{i=1}^m (\text{train}_i - y_i)^2}, \quad (16)$$

式中: m 为样本数据训练条目数量, train_i 是第 i 个样本数据的融合预测结果, y_i 为第 i 个样本数据的实测值。

Step 4: 计算每个粒子的适应度函数值。每个粒子个体需要记忆自身搜寻的最佳位置,第 i 个粒子自身搜寻过的最佳位置记作 $pbest_{id} = (p_{i1}, p_{i2}, \dots, p_{iD})$, 同时群体搜寻到的最佳位置也会被记录,记作 $gbest_d = (g_1, g_2, \dots, g_D)$ 。

Step 5: 根据公式 (8) 更新天牛左右须搜索空间位置,并分别计算 $f(\mathbf{x}_1)$ 和 $f(\mathbf{x}_c)$ 。

Step 6: 根据公式 (9) 计算搜索行为增量。

Step 7: 根据公式 (10) 公式 (11) 更新粒子速度和位置。

Step 8: 迭代控制。判断是否到达迭代次数或者当前位置的适应度函数值是否达到理想值,如果满足迭代停止条件,则进行下一步骤,否则继续从 Step 4 迭代搜索,并根据式 (12) ~ 式 (15) 分别更新惯性权重 w 、学习因子 c_1 和 c_2 以及天牛搜索行为增量的权重因子 λ 。

Step 9: 训练 BP 神经网络。种群最佳位置 $gbest_d$ 即是 BP 神经网络最佳初始权值和阈值,BP 神经网络利用 $gbest_d$ 作为初始参数进行训练学习直至权值和阈值参数确定。

3 仿真结果与分析

利用 MATLAB 2016a 进行仿真实验,与 GABP 算法和 PSOBP 算法进行对比实验,从数据融合精度、网络生命周期以及迭代收敛性能 3 个方面进行分析。无线传感网络仿真参数如表 1 所示。

为验证算法数据融合性能,本文采用森林火灾面积数据集为研究样本,该数据集提供了多项

表 1 仿真参数表

Table 1 Simulation parameters

参数	数值
传感区域	100 m×100 m
基站 (BS) 位置	(150 m,50 m)
节点数量 N	100
初始能量 E_0	0.5 J
功率放大能耗系数 ε_{fs}	10 pJ/bit·m ²
发射电路损耗 E_{elec}	50 nJ/bit
功率放大能耗系数 ε_{mp}	0.001 3 pJ/bit·m ⁴
簇首比例 p	5%
数据融合能耗 E_{DA}	5 nJ/bit·signal
数据大小	200 bits

气象数据包括温度、湿度、风速、降雨量等以及对应的森林火灾面积数据。无线传感器网络通过采集气象数据能有效监测小型火灾,对保障森林安全有重要意义。本文仿真实验中,BP 神经网络模型为 4 输入、1 输出的单隐藏层结构。选取数据集中传感器监测到的 4 项气象数据(温度、湿度、风速、降雨量)作为 BP 输入层参数,森林火灾面积数据作为输出层参数。隐藏层神经元的数量关系到网络的学习信息处理能力和结构复杂性,需要选择合适的隐藏层神经元数目。根据隐藏层神经元数量经验公式

$$k = \sqrt{m + n} + a, a = 1,2,3,\cdots,10, \quad (17)$$

隐藏层神经元数量取值范围为[4,12],表 2 为不同神经元数量下的融合误差 RMSE 值对比结果,其中 RMSE 量纲为森林火灾面积,单位为 hm²。

表 2 隐藏层神经元数量与 RMSE 值

Table 2 Neurons' number of hidden layers vs. RMSE

隐藏层神经元数量	误差 RMSE 值/hm ²
4	0.737 4
5	0.717 5
6	0.646 5
7	0.578 2
8	0.637 2
9	0.676 5
10	0.752 1
11	0.761 4
12	0.778 2

比较不同隐藏层神经元数量下的融合误差 RMSE 值,隐藏层神经元数量设置为 7 时,融合误差 RMSE 值最小,因此,本文中 BP 神经网络为 4-7-1 的单隐层结构,待优化参数个数为 4×7+7+7×1+1=43。

1)融合精度

融合精度对比采用平均相对误差 (average relative error, AvRE)、均方根误差 (root mean squared error, RMSE) 以及拟合优度 (goodness of fit, R^2) 作为评价指标,计算公式如下:

$$AvRE = \frac{\sum_{i=1}^n \frac{|u_i - y_i|}{y_i}}{N}, i = 1,2,\cdots,n, \quad (18)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^n (u_i - y_i)^2}, i = 1,2,\cdots,n, \quad (19)$$

$$R^2 = \frac{\sum_{i=1}^n (u_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, i = 1,2,\cdots,n. \quad (20)$$

式中: u_i 为第 i 个样本数据的融合值, y_i 为第 i 个样本数据的实际值, \bar{y} 为样本数据均值, n 为样本数据数量。理论上,平均相对误差和均方根误差两个评价指标越低,代表数据融合精度越高,拟合优度 R^2 取值范围为 0 至 1,越逼近 1 代表模型性能越佳。本文通过实验综合比较 BSO-BP 算法和 BP 算法、GA-BP 算法以及 PSO-BP 算法的融合精度,图 2 是 4 种算法的融合结果对比,由图可知,BSO-BP 算法的融合结果明显优于传统 BP 算法,融合结果和实际值差距较小,融合结果震荡较小,相比 GA-BP 算法和 PSO-BP 算法也有较大改善。

4 种算法的详细评价指标对比结果如表 3。从表 3 可以看出 BSO-BP 算法在平均相对误差、均方根误差以及拟合优度 3 个方面都优于其他 3 种算法。BSO-BP 算法和传统 BP 算法相比,平均相对误差和均方根误差分别减少 28% 和 37.4%。相较于 PSO-BP 算法,减少了 11% 的平均相对误差和 13.89% 的均方根误差。BSO-BP 算法克服了 BP 神经网络易陷入局部最优解的缺陷,增加了粒子群算法的种群搜索空间多样性,解决了粒子群算法早熟收敛的问题,因此 BSO-BP 算法相比其他算法有更好的数据融合效果。

2)迭代收敛对比

为测试 BSO-BP 算法的收敛性能,选择与 GA-BP 算法和 PSO-BP 算法进行仿真对比。图 3 为迭代收敛优化对比结果,由图可知,GA-BP 算法经过大约 55 次迭代收敛至最佳适应度值附近,PSO-BP 算法大约经过 46 次迭代能收敛至全局最优解附近,而 BSO-BP 算法只需大约 35 次迭代,

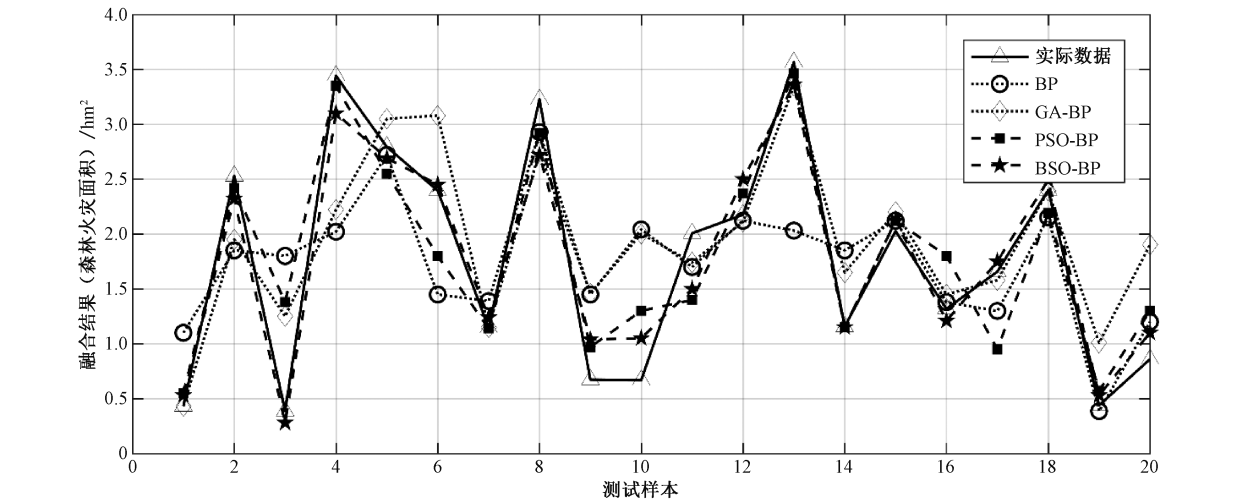


图 2 融合结果比较

Fig. 2 Comparison of fusion results

表 3 融合误差对比

Table 3 Comparison of fusion error

指标	BP	GA-BP	PSO-BP	BSO-BP
$AvRE$	0.460 3	0.440 7	0.372 5	0.331 4
$RmSE$	0.578 2	0.501 1	0.420 3	0.361 9
R^2	0.788 3	0.835 9	0.897 6	0.917 6

相比 GA-BP 和 PSO-BP 算法收敛速度提高,并且能寻找到更优的最佳适应度值。结果表明,将天牛须搜索算法中的天牛搜索行为增加到粒子群算法的粒子特性中能有效地增加寻优空间,避免陷入局部最优解,提高收敛速度和寻优性能,可以高效地优化 BP 神经网络的初始参数。

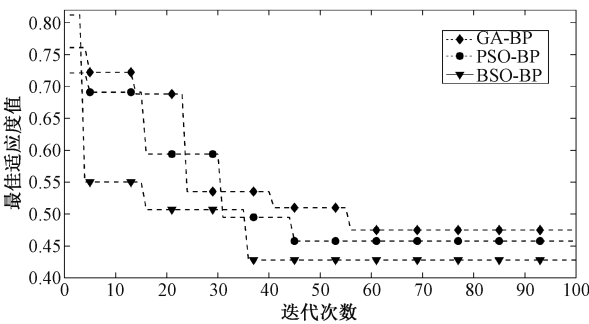


图 3 算法迭代收敛对比

Fig. 3 Comparison of iterative convergence among algorithms

3) 网络生命周期

WSN 中传感器节点能量有限,若监测区域出现单个或多个传感器节点能量耗尽或剩余能量不足以支持节点正常工作的情况,会导致监测盲区增多,获取的监测数据不准确,危险系数上升。网络生命周期在不同文献中有不同定义,文献[18]定义无线传感网络生命周期为目标监测区域内覆

盖率达到容忍下限值的时间,文献[19]定义网络生命周期为网络开始到节点能量全部消耗完的时间。本文采用文献[18]的定义,参考文献[11]和实际经验设定网络存活节点低于 85% 时,覆盖率达到容忍值,此时的数据可靠性下降,需要补充传感器节点能量。图 4(a)为各算法网络生命周期对比结果,可以看出在存活节点到达 85% 的阈值时,LEACH 协议网络运行至 700 轮左右,BSO-BP 算法和 BP 算法网络运行至 830 轮左右,相比 LEACH 协议延长 18.5%,这是由于 BSO-BP 算法和 BP 算法对监测数据进行了融合,减少了数据传输量,降低节点能耗,有效地延长了网络生命周期。图 4(b)为节点能耗对比,可以看出,LEACH 协议节点能耗明显高于 BSO-BP 算法和 BP 算法,而 BSO-BP 算法和 BP 算法的节点能耗和网络生命周期大致相当,这是因为两者的 BP 神经网络结构相同并且网络的构建和训练都是在基站节点进行,数据融合的输出维度一致。

4 结论

针对无线传感网 BP 神经网络数据融合算法存在对初始参数敏感、易陷入局部最优解的缺陷,本文提出基于改进粒子群的 BP 神经网络 WSN 数据融合算法(BSO-BP)。算法利用天牛须搜索(BAS)算法对粒子群算法进行改进,利用改进后的粒子群算法优化 BP 神经网络的初始权值和阈值,引入 WSN 数据融合中,簇首节点通过优化训练后的 BP 神经网络对采集数据进行特征提取,将融合后的数据发送至汇聚节点。与传统 BP 算

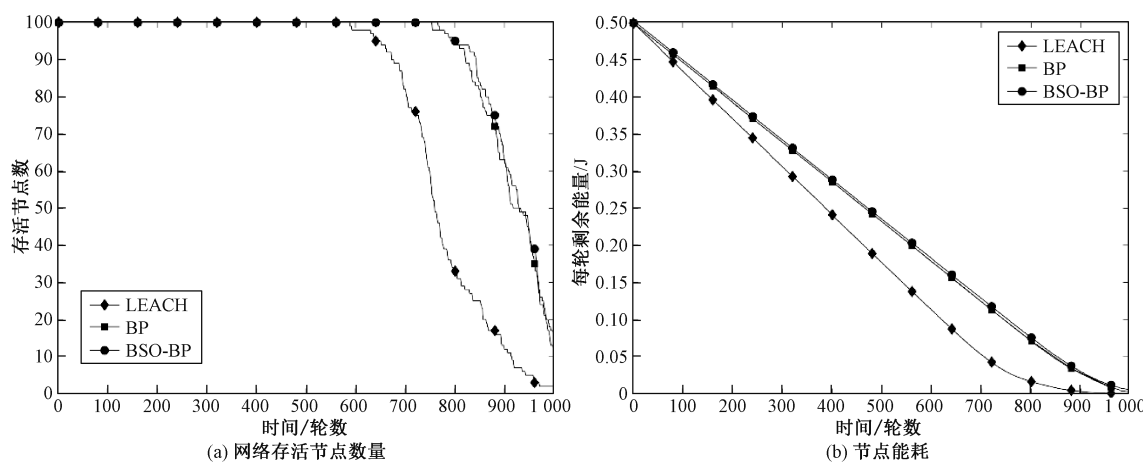


图 4 网络存活节点数量和节点能耗

Fig. 4 The number of alive nodes and the energy consumption of the node

法和 PSO-BP 算法进行对比,该算法能有效地提高融合精度和收敛速度,减少冗余数据传输,延长网络生命周期。

参考文献

[1] Lin D, Wang Q, Lin D, et al. An energy-efficient clustering routing protocol based on evolutionary game theory in wireless sensor networks [J]. International Journal of Distributed Sensor Networks, 2015, 2015: 1-12.

[2] Heinzelman W R, Chandrakasan A, Balakrishnan H. Energy-efficient communication protocol for wireless microsensor networks [C] // Proceedings of the 33rd Annual Hawaii International Conference on System Sciences. Hawaii: IEEE, 2000: 1-10.

[3] 孙凌逸, 黄先祥, 蔡伟, 等. 基于神经网络的无线传感器网络数据融合算法 [J]. 传感技术学报, 2011, 24 (1): 122-127.

[4] Boyinbode O, Le H, Mbogho A, et al. A survey on clustering algorithms for wireless sensor networks [C] // 2010 IEEE 13th International Conference on Network-based Information Systems. Takayama: IEEE, 2010: 358-364.

[5] 成瑜娟, 张敏洪. 感应半径可调的无线传感器网络生命周期的研究 [J]. 中国科学院研究生院学报, 2011, 28 (6): 768-775.

[6] 李兆祥. 无线传感器网络数据融合的算法研究 [D]. 武汉: 武汉理工大学, 2011.

[7] 赵亚光. 基于蚁群算法和 BP 神经网络的 WSN 数据融合算法研究 [D]. 昆明: 云南大学, 2013.

[8] 陈秋红, 郭猛. 基于 PSO-BP 的无线传感器网络数据融合算法研究 [J]. 计算机测量与控制, 2014, 22 (4): 1212-1214.

[9] Kennedy J, Eberhart R. Particle swarm optimization [C] //

Proceedings of ICNN'95-International Conference on Neural Networks. Perth: IEEE, 1995: 1942-1948.

[10] Wang T, Yang L, Liu Q. Beetle swarm optimization algorithm: theory and application [J]. arXiv preprint arXiv: 1808.00206, 2018.

[11] 胡强, 王海涛, 底楠, 等. 无线传感网中一种智能数据融合算法的实现及仿真分析 [J]. 传感技术学报, 2018, 31 (2): 283-288.

[12] Jiang X Y, Li S. BAS: beetle antennae search algorithm for optimization problems [J/OL]. International Journal of Robotics and Control, 2018 [2019-03-05]. <https://arxiv.org/pdf/1710.10724.pdf>.

[13] 刘杨, 田学锋, 詹志辉. 粒子群优化算法惯性权重控制方法的研究 [J]. 南京大学学报 (自然科学版), 2011, 47 (4): 364-371.

[14] 张震, 魏鹏, 李玉峰, 等. 改进粒子群联合禁忌搜索的特征选择算法 [J]. 通信学报, 2018, 39 (12): 60-68.

[15] 焦李成, 杨淑媛, 刘芳, 等. 神经网络七十年: 回顾与展望 [J]. 计算机学报, 2016, 39 (8): 1697-1716.

[16] 段大高, 谢永恒, 盖新新, 等. 基于神经网络的微博虚假信息识别模型 [J]. 信息安全, 2017, 17 (9): 134-137.

[17] 胡向东, 李秋实. 基于人工鱼群算法优化神经网络的 WSN 数据融合 [J]. 重庆邮电大学学报 (自然科学版), 2018, 30 (5): 614-619.

[18] Ishizuka M, Aida M. Performance study of node placement in sensor network [C] // 24th International Conference on Distributed Computing Systems. Hachioji: IEEE, 2004: 598-603.

[19] Guidoni D L, Boukerche A, Villas L A, et al. A tree-based approach to design heterogeneous sensor networks based on small world concepts [C] // 2011 IEEE 36th Conference on Local Computer Networks. Bonn: IEEE, 2011: 666-672.