

基于 RS-CSA-ELM 的 WSN 节点故障诊断^①



余正军

(汕头职业技术学院, 汕头 515078)

通讯作者: 余正军, E-mail: yuzhengjun123456@126.com

摘要: 为及时发现 WSN 节点故障隐患, 准确掌握 WSN 运行状态, 本文利用粗糙集理论属性约简算法 (简称 RS) 对 WSN 节点故障属性进行约简, 以最优的故障属性决策表重构训练样本数据集, 作为极限学习机 (Extreme Learning Machine, ELM) 神经网络的输入, 建立一个数据驱动的 WSN 节点故障断模型. 采用乌鸦搜索算法 (Crow Search Algorithm, CSA) 优化 ELM 神经网络的输入权值和隐含层阈值, 改善网络参数随机生成带来的 ELM 模型输出不稳定、分类精度偏低的问题. 通过对 RS-GA-ELM 模型进行仿真分析. 结果表明, RS-GA-ELM 模型能够在可靠性不同的数据集中, 保持较高的故障诊断效率, 符合 WSN 节点故障诊断的需求.

关键词: 故障诊断; 乌鸦搜索算法; 极限学习机; WSN; 粗糙集理论

引用格式: 余正军. 基于 RS-CSA-ELM 的 WSN 节点故障诊断. 计算机系统应用, 2021, 30(3): 221-226. <http://www.c-s-a.org.cn/1003-3254/7850.html>

Fault Diagnosis of WSN Nodes Based on RS-CSA-ELM

YU Zheng-Jun

(Shantou Polytechnic, Shantou 515078, China)

Abstract: In order to discover the hidden troubles of WSN nodes in time and accurately know the running status of WSN, this paper uses the attribute reduction algorithm of rough set theory (RS for short) to reduce the fault attributes of WSN nodes, and reconstructs the training sample data set with the optimal fault attribute decision table as an input to the Extreme Learning Machine (ELM) neural network. In this way, a data-driven fault diagnosis model of WSN nodes is established. The input weights and hidden layer thresholds of the ELM neural network are optimized through Crow Search Algorithm (CSA) to alleviate the unstable output and improve the low classification accuracy of the ELM model caused by the random generation of network parameters. Simulation analysis of the RS-GA-ELM model is carried out. The results show that the RS-GA-ELM model can keep efficiently diagnose faults in data sets with different reliability, which meets the needs of fault diagnosis of WSN nodes.

Key words: fault diagnosis; Crow Search Algorithm (CSA); Extreme Learning Machine (ELM); WSN; Rough Sets (RS) theory

随着智能物联网 (AIoT) 的兴起, 人们对智能化服务的需求也越来越多. 无线传感网 (WSN) 作为 AIoT 的关键组成部分, 其运行状态, 直接影响 AIoT 的服务质量. 然而, WSN 通常部署在环境恶劣的地方. 员工无法对 WSN 中的所有节点进行 24 小时实时观察和检查. 但是, 一旦 WSN 中节点出现故障, 将直接影响系统

最终的研判. 为此, 研究高效可靠的 WSN 节点故障诊断方法尤为重要, 它对延长 WSN 寿命, 保持系统稳定和提高系统服务质量具有重要的现实意义. 目前, WSN 节点故障诊断方法可分为基于统计学原理、故障分类和感知数据的时空特性分析三大类. 常用的诊断算法, 如文献 [1] 采用 2-回合投票故障诊断算法对故障进行

^① 收稿时间: 2020-07-16; 修改时间: 2020-08-13, 2020-09-01; 采用时间: 2020-09-08; csa 在线出版时间: 2021-03-03

诊断, 文献 [2] 以 ICA 进行故障属性的约简, 通过朴素贝叶斯进行故障诊断, 文献 [3] 利用 WSN 感知数据的时空相关性原理将诊断正常的候选簇头的观测值与簇内成员节点的观测值比较来对簇内成员节点进行诊断, 文献 [4] 以粗糙集进行故障属性的约简, 通过 GA-KELM 算法对故障诊断. 这些方法要么存在故障属性约简不彻底, 要么算法模型的参数复杂等问题. 本文鉴于他们的研究经验, 提出一种基于 RS-CSA-ELM 的 WSN 节点故障诊断方法, 以 RS 对故障属性进行约简, 并以 CSA 优化 ELM 参数作为故障诊断模型, 完成 WSN 节点故障诊断. 该算法仿真结果表明, 此法算法模型参数设置简单、易实现, 故障诊断的精确度高、误诊率低.

1 WSN 节点故障分析

1.1 WSN 网络结构

WSN 是一种高度自组织智能型数据处理网络, 通过节点的密集部署, 实现数据的量化收集、聚合和传输操作. 典型的 WSN 组成分为传感节点、中继节点、网关节点、调控中心 4 个部分. 其结构如图 1 所示.

1.2 WSN 节点结构及故障分析

WSN 节点自身就是一个高内聚低耦合的结构化产品, 它由感知单元、处理单元、信息收发单元及能量保障单元, 其结构如图 2 所示. 这 4 个独立的单元相互协作完成节点的任务, 任何一个单元出现出现故障将导致最终数据的错误影响 WSN 的正常工作. 而每个单元都具有自己的特征, 因此查找节点故障等同于查找节点产生故障的功能单元. 这种检测方法精度高, 便于维修^[5-7], 本文针对 WSN 网络结构及其节点结构, 将故障分为能量保障单元故障、处理单元故障、信息收发单元故障以及感知单元故障.

2 WSN 节点故障诊断算法设计

2.1 故障诊断属性约简算法

粗糙集理论属性约简算法 (简称 RS) 能从模糊数据本身归纳出数据关键信息, 删除冗余信息, 求得最小表达式, 建立决策表. 正好可用于对 WSN 节点故障诊断属性的约简, 其过程如下: 设故障诊断决策表为 $S = (U, R, V, f)$, 其中, S 为 WSN 节点故障诊断决策表; $R = P \cup D$ 节点故障属性集合, P, D 分别代表条件属性集合和决策属性集合; $U = \{x_1, x_2, \dots, x_m\}$ 为论域; 属性

值域 $V = \cup V_R(m[i]) \in R$; 则信息函数为 $f: U \times R \rightarrow V$, 表示对每个 $m \in R, x \in U, f(x, m) \in V_R$. 其分辨矩阵可设置成:

$$C_D(i, j) = \begin{cases} m[k], & d(x_i) = d(x_j) \\ 0, & d(x_i) \neq d(x_j) \end{cases}$$

其中, $C_D(i, j)$ 为分辨矩阵的第 i 行第 j 列处的元素, $m[k] \in P$.

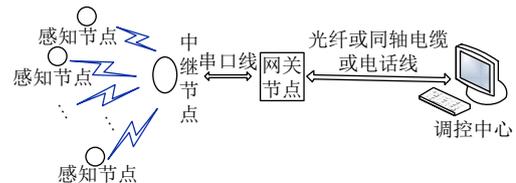


图 1 典型 WSN 网络结构

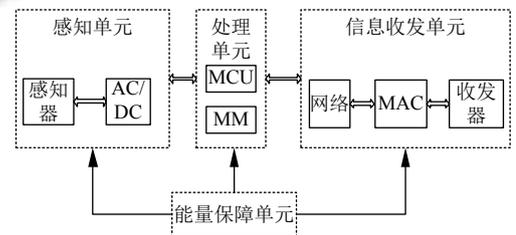


图 2 WSN 节点结构

基于上述分辨矩阵的故障属性约简算法步骤如下:

- (1) 利用上文描述, 根据故障样本决策表构建分辨矩阵 $C_D(i, j)$.
- (2) 约简分辨矩阵 $C_D(i, j)$, 将其中的单属性元素对应位置零.
- (3) 若 $C_D(i, j) \neq 0$, 且 $C_D \neq \emptyset$ 时, 则构建 RS 属性样本 T 约简的操作范式:

$$T = \bigwedge_{C_D(i, j) \neq 0} \{ \bigvee_{m[k] \in C_D(i, j)} m[k] \}$$

- (4) 向步骤 (3) 的 RS 属性样本 T 中每一项都添加步骤 (2) 中的单属性, RS 属性样本约简 T' , 其每项都代表一个故障属性约简.

2.2 WSN 节点故障诊断模型

RS-CSA-ELM 算法对故障诊断原理描述为: 以 WSN 节点故障预兆构建故障样本数据集, 以故障种类属性和特征属性构建故障决策表. 首先, 根据 RS 剔除原始故障样本属性冗余信息, 获得约简后的故障属性组合; 其次, 利用属性之间的相关性选取最优属性约简; 以最优属性约简构建新的故障样本作为算法模型的训练样本, 最后, 建立 RS-CSA-ELM 故障诊断模型对 WSN 节点故障进行诊断, 如图 3 所示.

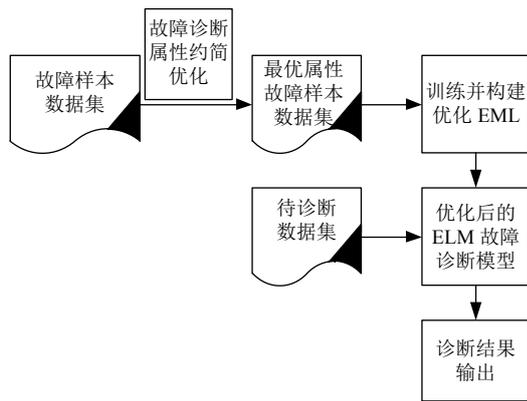


图3 WSN节点的RS-CSA-ELM故障诊断模型

2.3 CSA-ELM 算法

(1) 极限学习机 (ELM) 原理

设给定 N 组 WSM 节点数据样本 $(x_i, t_i) \in R^n$, 对于输入层、隐含层和输出层的节点数分别为 n, L, m 的 ELM 的学习机, 若采用 Sigmoid 类型的激励函数 $G(x)$, 则 ELM 隐含层输出即为:

$$f_L(x) = \sum_{i=1}^L \beta_i G(\alpha_i \cdot x_j + \gamma_i) = t_i, j = 1, 2, \dots, N$$

其中, $\beta_i = [\beta_{i1}, \beta_{i2}, \dots, \beta_{im}]^T$, $\alpha_i = [\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{in}]^T$, $\gamma_i = [\gamma_{i1}, \gamma_{i2}, \dots, \gamma_{im}]^T$; 隐含层的第 i 个神经元与输出层的连接权值为 β_i ; 与输入层的连接权值 α_i ; 阈值 γ_i [8]. 则 $f_L(x)$ 矩阵表示形式如下:

$$H\beta = T$$

其中,

$$H = \begin{bmatrix} G(\alpha_1 \cdot x_1 + \gamma_1) & \cdots & G(\alpha_L \cdot x_j + \gamma_L) \\ \vdots & \ddots & \vdots \\ G(\alpha_1 \cdot x_N + \gamma_1) & \cdots & G(\alpha_L \cdot x_N + \gamma_L) \end{bmatrix}_{N \times L}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{bmatrix}_{L \times m}, T = \begin{bmatrix} t_1^T \\ \vdots \\ t_L^T \end{bmatrix}_{N \times m}$$

当 $G(x)$ 无限可微时, 参数 α_i 与 γ_i 随机产生, 且在训练过程中不需要被调整, ELM 模型训练过程可看作求解线性系统 $H\beta = T$ 关于 β 的最小二乘解, 即 $\beta = H^+T$, 其中 H^+ 为 H 的 Moore-Penrose 广义逆.

(2) CSA 优化 ELM

极限学习机 (ELM) 是一类基于前馈神经网络的训练算法, 其具有学习速度快、泛化性能强、不易陷入局部极值的特性, 因此被广泛应用 [9]; 但由于该算法中

输入权值 α_i 和隐含层阈值 γ_i 是随机选取导致其隐含层神经元不具备自我调节能力. 针对此缺陷本文采用 CSA 算法对 ELM 输入权值 α_i 和隐含层阈值 γ_i 进行优化, 构建 CSA-ELM 神经网络, 提高模型诊断的精确度和稳定性 [10].

乌鸦搜索算法 (CSA) 是仿效乌鸦在储备食物和盗窃食物的智能行为而形成的一种基于种群的优化算法, CSA 与其它神经网络参数优化算法如 GA、POS 一样, 均是利用种群的优势来增加寻找最优解的概率, 同时避免陷入局部最优解. 但优化算法的参数设置对其性能影响很大, 所以参数越少的算法越易应用. 种群优化算法除种群大小和最大迭代次数外, 还要设置其它参数, 如 GA 还需再设置选择方法、交叉方法、交叉概率、变异方法以及变异概率; PSO 还需再要设置有惯性权重、粒子自身的学习因子和社会学习因子; 而只需再设置乌鸦的飞行长度和辨识概率即可 [11-13].

另外, CSA 不是贪心算法, 如在搜索空间范围内只要没有比乌鸦产生的新位置更优, 它依然会飞向新位置, 这样产生的解是多样的, 使得算法具有最大概率寻找最优解功能. 乌鸦跟踪行为是随机的, 这样就增大了算法搜索空间, 即 CSA 具有全局搜索能力. 算法在迭代过程中, 种群中所有乌鸦都在不断地更新藏食地点, 从而保证种群中所有的藏食地点是最优的, 并利用乌鸦记忆记录最优存食位置. 迭代结束后, 找到种群中乌鸦的适应度值最优的记忆作为最后求到的最优解. 表 1 为 CSA 与 GA 和 PSO 算法的性能比较.

表1 CSA 与 GA 和 PSO 算法性能比较

算法	特性	缺陷
CSA	全局寻优、收敛速度快、精度高、参数少、易实现	循环开销大
GA	全局寻优、内在并行、分布式计算	易陷入局部最优、收敛速度慢、参数多、编码和解码操作复杂、工作量大
PSO	早期收敛速度快、寻优精度高	后期收敛速度慢、易陷入局部最优、参数多

CSA 算法应用到 ELM 的输入权值 α_i 和隐含层阈值 γ_i 选取, 构建 CSA-ELM 算法. 在该算法中, 将 ELM 训练数据的输入权值 α_i 和隐含层阈值 γ_i 映射为 CSA 种群中乌鸦所处的每个位置和藏食物的记忆, CSA 的适应度值对应于 ELM 的训练误差, 将求取最优输入权值、阈值问题转化为计算 CSA 适应度值即选择最优

乌鸦记忆问题. 图4给出 CSA-ELM 算法流程, 主要包括 ELM 网络确定、CSA 算法优化和 ELM 网络的训练、预测等.

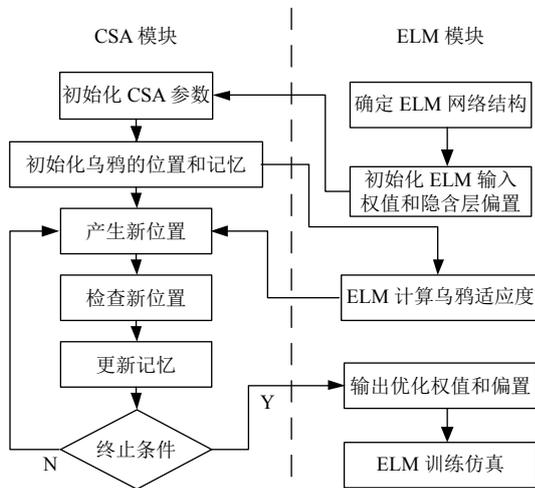


图4 CSA-ELM 算法流程

该算法的详细操作要领如下:

Step 1. 算法初始化. 确定 ELM 各层的神经元个数, 随机生成 ELM 的输入权值 α_i 和隐含层阈值 γ_i . CSA 的种群搜索空间设置成二维, 并将 α_i 和 γ_i 作为乌鸦初始所处位置和记忆即公式表示为: $x^{i,iter} = (\alpha_i, \gamma_i)$ 、 $m^{i,iter} = (\alpha_i, \gamma_i)$, 对 CSA 进行初始化.

Step 2. 计算乌鸦适应度. 适应度计算公式为:

$$f(\cdot) = \frac{N_e}{N}$$

其中, N_e 为正确的样本数, N 为训练样本总数.

Step 3. 更新乌鸦记忆. 更新的计算公式为:

$$m^{i,iter+1} = \begin{cases} x^{i,iter+1}, & f(x^{i,iter+1}) \text{ is better than } f(x^{i,iter}) \\ m^{i,iter}, & o \cdot w \end{cases}$$

检查终止条件, 若满足则输出最优的 α_i 和 γ_i , 否则转向 Step 2.

Step 4. ELM 训练和诊断. 将最终乌鸦位置和记忆作为最优的 ELM 的输入权值和阈值赋予 ELM, 使用训练样本对 ELM 进行训练, 采用公式 $\beta = H^+T$ 计算输出层权值. 最后将待诊样本注入到 ELM 模型进行诊断输出.

2.4 RS-CSA-ELM 算法步骤

根据上文所述, 将 RS 的故障属性约简算法与经 CSA 优化的极限学习机有机结合, 构建基于 RS-CSA-ELM 的 WSN 节点故障诊断算法, 其过程如下:

Step 1. 根据 WSN 节点组成结构, 构建符合实际情况的 WSN 节点故障属性表并做量化处理.

Step 2. 根据上文所提的故障属性约简算法对 Step 1 构建的 WSN 节点故障属性表进行属性约简, 删除冗余的故障属性信息, 得到最优的故障诊断决策表.

Step 3. 根据 Step 2 得到的故障属性决策表重构训练样本数据集.

Step 4. 将 Step 3 得到的训练样本数据集作为 RS-CSA-ELM 算法模型的输入, 利用 CSA 优化 ELM 的参数, 即利用训练的故障样本数据对 ELM 进行训练, 使其达到最佳的 ELM 故障模型, 而后用该模型对待诊样本数据进行故障诊断, 实现 WSN 节点故障任务.

3 仿真与结果分析

3.1 仿真实验设置

(1) WSN 节点故障属性的约简

为验证上述算法得有效性, 先构建其故障属性及量化表(表2), 再将表2转换为故障诊断决策表(表3). 最后采前文中设计的 RS 故障属性约简算法对表2进行约简, 以获得最优故障诊断决策.

表2 部分故障属性及量化处理表

编号	故障预兆	量化处理
p[1]	待诊节点对询问是否有响应?	是: p[1]=0; 否: p[1]=1
p[2]	待诊节点是否定期有信号返回?	是: p[2]=0; 否: p[2]=1
p[3]	待诊节点是否正确执行指令?	是: p[3]=0; 否: p[3]=1
p[4]	待诊节点是否能传递其他节点的数据?	是: p[4]=0; 否: p[4]=1
p[5]	待诊节点是否有感知数据?	是: p[5]=0; 否: p[5]=1
p[6]	待诊节点数据是否严重偏离相邻节点数据?	是: p[6]=1; 否: p[6]=0
p[7]	待诊节点监测值是否偏离平均值?	是: p[7]=1; 否: p[7]=0
p[8]	待诊节点发射频率改变, 是否有响应?	是: p[8]=1; 否: p[8]=0

表3 部分故障诊断决策表

编号	故障种类	相应故障预兆
d_1	正常	无
d_2	处理单元	p[1] p[2] p[3] p[4]
d_3	感知单元	p[5] p[6] p[7]
d_4	信息收发单元	p[1] p[2] p[3] p[4] p[8]
d_5	能量保障单元	p[1] p[2] p[3] p[4] p[5] p[6] p[7] p[8]

(2) 算法参数设置

仿真设置算法初始条件: 飞行长度 $f_l = 0.2$, 最大迭代次数 $t_{max} = 50$, 意识概率 $AP = 0.1$, ELM 的输入权值 $\alpha_i = rand(0, 1)$, 隐含层阈值 $\gamma_i = rand(0, 1)$, WSN 节点数量作为乌鸦只数 $N=200$.

为验证所提算法的性能及可行性,在 Ubuntu16.04 环境下用 Python3.5+ Tensorflow 编程仿真,同时考虑到 WSN 实际工作环境,实际所收集的数据参与真实数据之间存在误差.模拟在同一条件下对优化前后 EML 性能进行仿真试验,同时也对本文所提算法与最新文献所提算法作了对比仿真实验.

3.2 仿真结果分析

首先,在相同条件下,对 CSA 优化前的 ELM 与优化后的 ELM 模型进行了对比测试,其结果如图 5、图 6 所示.

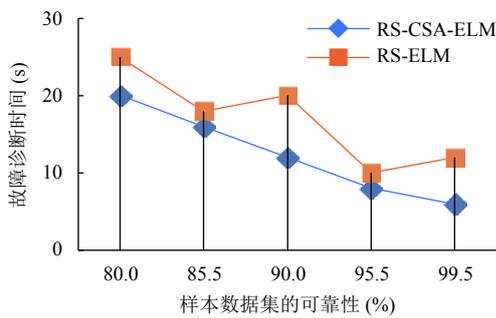


图 5 CSA 优化前后的 ELM 故障诊断用时对比图

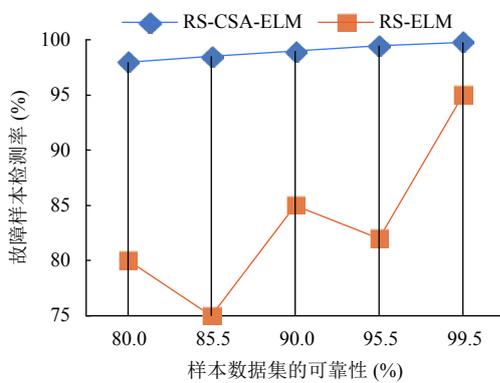


图 6 CSA 优化前后的 ELM 故障检测精度比较图

其次,对 CSA 优化 RS-CSA-ELM 模型的参数的做了仿真验证.其结果如图 7 所示.

最后,对可靠性为 80%、85.5%、90% 和 99.5% 样本数据集进行约简后,将本文提的 RS-CSA-ELM 算法与文献 [14] 的 RSOPNN 算法及文献 [15] 的 RS-PSO-KELM 算法进行比较测试.其结果如图 8 及图 9 所示.

从图 5、图 6 可以看出经过 CSA 优化之后的 ELM 其稳定性得到了很大的提高;从图 7 可以看出 CSA 在迭代到 15 次的时即可寻找到最优得 ELM 输入权值和隐含层阈值;从图 8 可以看出 3 种算法对 WSN

节点故障诊断用时都是随样本数据集可靠性增加而减少,但本文所提的 RS-CSA-ELM 算法与其它两种算法在相同条件下用时是最短的;从图 9 可看出 3 种算法的诊断准确性随样本数据集可靠性增加而提高,但本文所提的 RS-CSA-ELM 算法与其它两种算法在相同条件下诊断的准确性是最高的.综上所述,本文所提算法能够快而准的对 WSN 节点故障进行诊断识别,且稳定性和诊断的准确性可满足实际应用要求.

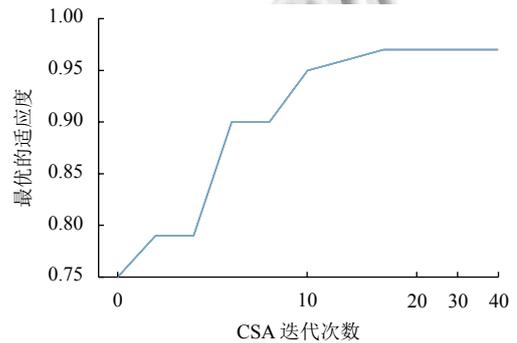


图 7 CSA 对 RS-CSA-ELM 模型寻优迭代次数图

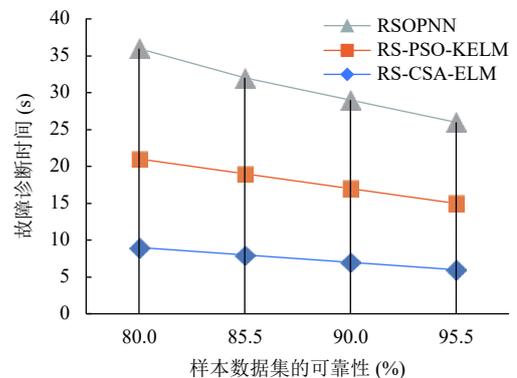


图 8 算法诊断用时比较图

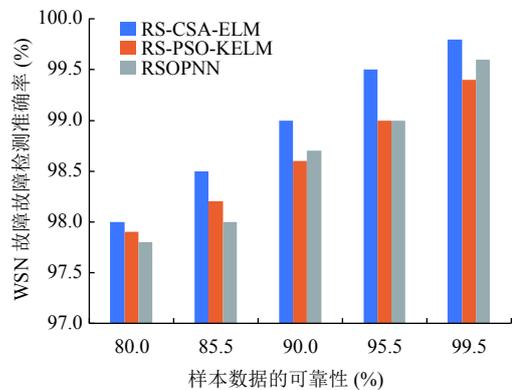


图 9 算法性能比较图

4 结束语

鉴于现有 WSN 节点在故障诊断方法的缺陷^[16], 提出 RS-CSA-KELM 算法予以解决, 通过粗糙集理论属性约简算法 (简称 RS) 对故障属性约简、利用 CSA 算法对 EML 的参数进行优化, 建立数据驱动式的 RS-CSA-ELM 故障诊断模型. 该模型对 WSN 节点故障的诊断快而准, 并能有效的解决误诊和漏诊等问题, 特别适合用于拓扑结构实时变化的网络的故障诊断. 当然, 所提之法仍需研究、改进, 如进一步考虑算法中各参数设置及各算法之间的优缺点互补等问题.

参考文献

- 1 丁晗. 无线传感器网络的故障节点检测算法设计与分析 [硕士学位论文]. 苏州: 苏州大学, 2016.
- 2 张新华. 基于 ICA 独立成分和加权依赖贝叶斯的传感器节点故障诊断. 重庆师范大学学报 (自然科学版), 2015, 32(2): 138-142.
- 3 曾庆杰, 朱晓娟, 王先超. 基于节能的 WSN 分簇式故障诊断算法. 阜阳师范学院学报 (自然科学版), 2019, 36(4): 54-60.
- 4 刘建伟. WSN 节点故障诊断方法研究及测试平台设计 [硕士学位论文]. 长沙: 湖南大学, 2017.
- 5 Sun YM, Liu XJ, Chen XG, *et al.* Research and application of a fault self-diagnosis method for roots flowmeter based on WSN node. *Wireless Personal Communications*, 2017, 95(3): 2315-2330. [doi: 10.1007/s11277-017-4104-8]
- 6 Duh DR, Li SP, Cheng VW. Distributed fault-tolerant event region detection of wireless sensor networks. *International Journal of Distributed Sensor Networks*, 2013, 9(9): 286-291.
- 7 Oh SH, Hong CO, Choi YH. A malicious and malfunctioning node detection scheme for wireless sensor networks. *Wireless Sensor Network*, 2012, 4(3): 84-90. [doi: 10.4236/wsn.2012.43012]
- 8 周奚, 薛善良. 基于改进的粗糙集和神经网络的 WSN 故障诊断. *计算机科学*, 2016, 43(S2): 21-25.
- 9 郑茂辉, 刘少非. GA 优化 ELM 神经网络的排水管道缺陷诊断. *哈尔滨工业大学学报*: 1-7. <http://kns.cnki.net/kcms/detail/23.1235.T.20200420.1814.002.html>, [2020-04-29].
- 10 蒋鹏. 一种改进的 DFD 无线传感器网络节点故障诊断算法研究. *传感技术学报*, 2008, 21(8): 1417-1421. [doi: 10.3969/j.issn.1004-1699.2008.08.026]
- 11 林涛, 郝章肖, 冯竞凯. 基于改进乌鸦搜索算法的云计算任务调度研究. *微电子学与计算机*, 2020, 37(2): 20-24.
- 12 王丽婷, 张金鑫, 张金华. 乌鸦搜索算法在 SVM 参数优化中的应用. *计算机工程与应用*, 2019, 55(21): 214-219. [doi: 10.3778/j.issn.1002-8331.1903-0381]
- 13 Sun QY, Sun YM, Liu XJ, *et al.* Study on fault diagnosis algorithm in WSN nodes based on RPCA model and SVDD for multi-class classification. *Cluster Computing*, 2019, 22(3): 6043-6057.
- 14 李洋, 高岭, 孙骞, 等. 基于 RSOPNN 的无线传感器网络节点故障诊断算法. *计算机工程与应用*, 2017, 53(9): 111-116, 157. [doi: 10.3778/j.issn.1002-8331.1511-0185]
- 15 柴继超. 无线传感器网络节点故障诊断方法研究 [硕士学位论文]. 哈尔滨: 哈尔滨理工大学, 2019.
- 16 Xing XD. Research on fault diagnosis technology of wireless sensor network based on improved frog leaping algorithm. *Machine Tool & Hydraulics*, 2019, 47(12): 97-102. [doi: 10.3969/j.issn.1001-3881.2019.12.015]