

# 基于自适应广义回归神经网络的链路质量评估

舒 坚 高 素 陈宇斌

(南昌航空大学软件学院 南昌 330063)

(shujian@nchu.edu.cn)

## Link Quality Estimator Based on Adaptive General Regression Neural Network

Shu Jian, Gao Su, and Chen Yubin

(School of Software, Nanchang Hangkong University, Nanchang 330063)

**Abstract** How to appropriately select link quality metrics and build a link quality estimation model with better performance, generalization capability, and lower time complexity is one of the key challenges in wireless sensor networks. We select  $M_{CS} = \{\mu, r, \sigma^2\}$ , where  $\mu = \{\mu_{lqi}, \mu_{rssi}, \mu_{snr}\}$ ,  $r = \{r_{lqi}, r_{rssi}, r_{snr}\}$ ,  $\sigma^2 = \{\sigma_{lqi}^2, \sigma_{rssi}^2, \sigma_{snr}^2\}$ , as link quality metric candidate set. A link quality metric selection algorithm based on wrapper is proposed, which employs adaptive general regression neural network (AGRNN) to evaluate the importance of link quality metric candidate subsets so as to select link quality metrics. Taking advantages of general regression neural network (GRNN) in classification and time complexity, this paper proposes a link quality estimation model based on AGRNN which assigns different smoothing factors to each link quality metric and adaptively modifies them by using back propagation. Indexes, such as accuracy, recall, generalization error and computing time, are used to evaluate the link quality estimation models. In scenarios of the indoor scenario, the park scenario and the motorway scenario, the experimental results show that the proposed link quality estimation model can achieve better performance, generalization capability, and lower time complexity, compared with link quality estimation models based on polynomial, support vector classifier and random forest respectively.

**Key words** wireless sensor networks; link quality estimation; metric selection algorithm based on wrapper; adaptive general regression neural network

**摘 要** 为选择合适的链路质量参数,进一步提高链路质量评估的性能和泛化能力、降低时间复杂度,确定链路质量参数的备选集  $M_{CS} = \{\mu, r, \sigma^2\}$ , 其中  $\mu = \{\mu_{lqi}, \mu_{rssi}, \mu_{snr}\}$ ,  $r = \{r_{lqi}, r_{rssi}, r_{snr}\}$ ,  $\sigma^2 = \{\sigma_{lqi}^2, \sigma_{rssi}^2, \sigma_{snr}^2\}$ ; 提出包裹式链路质量参数选取算法,采用自适应广义回归神经网络(adaptive general regression neural network, AGRNN)评价各备选子集的重要性,选择链路质量参数;借助广义回归神经网络(general regression neural network, GRNN)在分类以及时间上的优势,提出基于 AGRNN 的链路质量评估模型,该模型为每个链路质量参数分配不同的光滑因子,采用误差反向传播的思想对其进行自适应修正;采用准确率、召回率、泛化误差和计算时间评价链路质量评估模型.室内、公园和公路场景下的

收稿日期:2019-08-23;修回日期:2020-02-19

基金项目:国家自然科学基金项目(61962037,61762065);江西省自然科学基金重点项目(20202BABL202039,20181BAB202015);江西省研究生创新专项资金项目(YC2018-S384)

This work was supported by the National Natural Science Foundation of China (61962037, 61762065), the key program of the Natural Science Foundation of Jiangxi Province (20202BABL202039, 20181BAB202015), and the Innovation Foundation for Postgraduate Student of Jiangxi Province (YC2018-S384).

实验表明:与基于多项式法、随机森林、支持向量分类器的链路质量评估模型相比,基于 AGRNN 的链路质量评估模型具有更优的评估性能和更好的泛化能力以及更低的时间复杂度。

**关键词** 无线传感器网络;链路质量评估;包裹式参数选取算法;自适应广义回归神经网络

**中图分类号** TP393.17

无线传感器网络(wireless sensor networks, WSNs)是物联网的重要组成部分,用于获取有关环境的条件和个人信息,充当物理世界和数字世界间的重要桥梁.WSNs 由大量传感器节点组成,其中传感器节点可控制或监测物理信息及环境条件,例如声音、振动、压力、温度和污染物的运动,并通过网络协同技术传递到基站.WSNs 广泛应用于与环境监测、医疗保健、军事、国家安全等相关领域中<sup>[1]</sup>.

传感器节点大多采用无线射频进行通信,故 WSNs 的通信质量易受环境的影响,例如正在使用某一范围的频谱、特定调制方案,以及通信设备本身和随机信号的干扰.随着时间的推移,通信质量会发生变化,节点之间的链路是不可靠的<sup>[2]</sup>.针对链路质量评估问题中存在的 2 个挑战:1) 如何选择合适的链路质量参数.2) 如何构建具有较优评估性能、较好泛化能力和较低时间复杂度的链路质量评估模型.本文提出基于自适应广义回归神经网络的链路质量模型.

本文的主要贡献包括 2 个方面:

1) 提出基于包裹式的链路质量参数选取算法,采用自适应广义回归神经网络(adaptive general regression neural network, AGRNN)评价各备选子集的重要性,选取链路质量参数.

2) 提出基于 AGRNN 的链路质量评估模型,该评估模型可为每个链路质量参数分配不同的光滑因子并进行自适应调整,且具有更低的时间复杂度,能更有效地评估链路质量.

## 1 相关研究

目前,国内外学者对 WSNs 链路质量评估方法进行了探索,主要包括:基于硬度量、软度量的链路质量评估方法;基于混合多层的综合性链路质量评估方法;基于映射关系的链路质量评估方法以及基于数据驱动的链路质量评估方法<sup>[3]</sup>.

1) 基于硬度量、软度量的链路质量评估方法

基于硬度量链路的质量评估方法主要采用接收信号强度指示(received signal strength indicator,

RSSI)、链路质量指示(link quality indicator, LQI)、信噪比(signal to noise ratio, SNR)等物理层参数对链路质量进行评估<sup>[2]</sup>.文献[4]提出一种快速判断链路质量的评估方法,当  $lqi \leq 165$  时认为链路质量较差,当  $lqi = 255$  时认为链路质量较好,实验表明:与基于 RSSI 的链路质量评估模型相比,基于 LQI 的链路质量评估模型能给出更加准确的评估结果,且稳定性相对较好.

基于软度量链路的质量评估方法主要采用包接收率(packet reception ratio, PRR)、需要传输次数(required number of packet retransmissions, RNP)、期望传输次数(expected transmission count, ETX)等统计量对链路质量进行评估.文献[5]提出一种基于增强卡尔曼滤波器的链路质量评估方法,根据每个分组的接收 I/Q 流的观察来估计链路质量,通过卡尔曼滤波过滤掉测量噪声,在工业现场的环境验证了基于增强卡尔曼滤波器的链路质量评估方法的有效性.

基于硬度量的链路质量评估方法可直接从硬件中读取,消耗能耗较少,容易受噪声、多径衰落和干扰等影响,对链路质量进行评估具有一定的敏捷性;基于软度量的评估方法能够较准确地对链路质量进行评估,反映链路的稳定性,但需发送大量的探测包,统计结果得到,该方法会消耗较多的能耗.故可知单独的硬度量或者软度量均难以较全面的评估链路质量.

2) 基于混合多层的综合性链路质量评估方法

基于混合多层的综合性链路质量评估方法通过混合链路的物理层、链路层等多层信息对链路质量进行评估.文献[6]提出了一种链路质量评估模型——F-LQE(fuzzy link quality estimator),该模型基于分组传送、非对称性、稳定性和信道质量 4 种链路质量属性对链路质量进行评估,采用模糊逻辑对各链路属性进行融合.实验结果表明:基于 F-LQE 的链路质量评估模型具有较好的稳定性及可靠性;文献[7]提出一种混合的链路质量评估度量(link verification metric, LVM),将 3 个链路质量参数进行融合:

$$lvm = prr \times \frac{Mapped(snr)}{1 - Mapped(rssi)}, \quad (1)$$

其中, *Mapped* 表示标准化函数, 实验结果表明 LVM 可借助硬件参数 SNR 的敏捷性优势较为准确地对具有相似 *pr*r 值的稳定链路进行区分, 对于具有相似 *snr* 和 *rss*i 值的不稳定链路, LVM 主要依赖于 PRR 表征链路质量, LVM 是一种低成本的链路质量评估度量, 未来可在路由决策上扮演重要角色, 但是 LVM 的有效性仅在软件环境下被验证。

基于混合多层的综合性链路质量评估方法混合了多层的链路信息, 能较准确地评估链路质量, 但是需要同时收集多层信息, 链路质量评估的及时性得不到很好保证。

### 3) 基于映射关系的链路质量评估方法

硬度量对环境变化敏感且耗能低, 而软度量具有一定的稳定性, 在此基础上寻找物理层参数与 PRR 等链路性能参数之间的直接映射关系得到了广泛的研究: 文献[8]提出了一种三角链路质量度量法和基于最小路径间干扰的地理多径路由协议, 三角链路质量度量法是一种混合的链路质量度量方法, 它将 *snr*, *lqi*, *pr*r 几何地合并为一个稳健的估计器, 其中 *snr* 和 *lqi* 的平均值作为直角边, 斜边的值为度量值, 三角链路质量度量法提供了必要的链路质量特性, 可供路由选择可靠的转发节点, 实验表明: 应用三角链路质量度量的协议可以在合理的能源成本下改善网络寿命, 确保高数据包传输率。

基于映射关系的链路质量评估方法综合考虑了链路质量评估的稳定性、敏捷性以及能耗等特性, 但是融合方法过于简单, 评估动态不稳定链路的质量时准确性较差。

### 4) 基于数据驱动的链路质量评估

基于数据驱动的链路质量评估方法是以数据为基础, 经过训练以及逼近真实值等过程形成自动化的链路质量评估模型, 包括文献[9]提出基于多项式法 (polynomial, Poly) 的链路质量评估模型, 仅以 *rss*i 的均值作为输入, 仅当多项式的最高次幂为 6 时评估性能最佳, 在工厂以及室外环境下验证了该评估模型的有效性; 文献[10]采用机器学习方法结合 WSNs 环境属性, 例如探测到的节点个数、成功接收数据包个数、数据包间间隔等, 对链路质量进行评估, 实验结果表明: 与回归树、线性回归相比, 基于人工神经网络 (artificial neural network, ANN) 的评估结果更加接近真实值; 文献[11]采用小波神经网络得到下一时刻的 SNR 及其方差, 利用 SNR 与 PRR 之间的映射关系, 计算出 PRR 的置信区间, 在智能电网环境下验证了该方法的有效性; 文献[12]

提出了一种基于支持向量分类器 (support vector classifier, SVC) 的链路质量评估算法, 选取 RSSI 和 LQI 作为链路度量, 通过实验法确定径向基函数 (radial basis function, RBF), 构建基于 SVC 的链路质量评估模型, 实验表明基于 SVC 的链路质量评估模型的准确率为 93.38%, 优于 LQI-PRR 和 FLI 这 2 种算法; 文献[13]选取一些最具代表性的机器学习算法分别为: 逻辑回归、随机森林 (random forest, RF)、支持向量分类器、多层感知机、决策树和多数算法, 对链路质量进行评估, 实验结果表明基于 RF 的链路质量评估模型的准确率为 95.3% 并优于其他链路质量评估模型。

基于数据驱动的链路质量评估方法, 利用智能的学习算法对链路质量参数与链路质量之间的映射关系进行学习, 保证了评估的准确性, 但是评估模型的时间复杂度会随着样本的增加而急剧增加。

广义回归神经网络 (general regression neural network, GRNN) 是由 Specht<sup>[14]</sup> 在 1991 年提出的。GRNN 建立在非参数回归的基础上, 以样本数据为后验条件, 执行 Parzen 非参数估计, 依据最大概率原计算网络输出。由于 GRNN 具有训练速度快、非线性映射能力强、容错性高、鲁棒性强等特点, 已被广泛应用包括噪声数据处理、函数逼近、主动振动控制以及故障诊断等回归以及分类问题<sup>[15-16]</sup>。由于 WSNs 易受外部环境的噪声、电磁波以及多径效应的影响, GRNN 中单一的光滑因子难以适应这种复杂的环境, 而 AGRNN 可为每个输入参数分配不同的光滑因子并进行自适应调整, 在环境嘈杂的实验中进行测试, 结果表明 AGRNN 具有良好的建模性能<sup>[17]</sup>。

基于映射关系的链路质量评估方法综合考虑链路质量评估的特性, 本文借鉴该方法选取物理层参数为链路质量参数, 根据 PRR 划分链路质量等级作为标签, 构建链路质量参数与链路质量等级的关系模型; 在基于数据驱动的链路质量评估方法中, 人工神经网络、支持向量分类器、随机森林、小波神经网络运用于链路质量评估和预测领域取得了较好的效果, AGRNN 是一种具有快速学习和良好建模能力的人工神经网络, 故本文采用时间复杂度较低的 AGRNN 构建链路质量评估模型。

## 2 链路质量等级的划分

链路质量等级的划分主要分为硬划分和软划分。软划分一般采用聚类算法对样本进行聚类, 根据聚簇中 *pr*r 的范围确定划分链路质量等级的阈值,

样本的变化可能会导致阈值发生变化;硬划分一般借助如平稳概率、运行或损失长度等物理参数来区分链路质量的主要类别,各类别的链路质量分别对应一个  $prrr$  的范围,即划分的阈值,阈值不受样本变化的影响.文献[18]应用 Gilbert-Elliot 模型对  $PRR$  进行分析,借助物理参数将  $prrr \geq 0.8$  的链路划分为“好”链路,将  $prrr < 0.2$  的链路划分为“差”链路,余下为“中等”链路,在此基础上文献[19]采用硬划分方法划分链路质量等级,评估模型具有较好的准确性.硬划分方法简单,且阈值设定合理,故本文借鉴文献[18]划分链路质量等级,如图 1 所示:

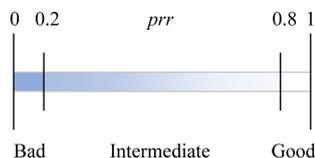


Fig. 1 Link quality levels divided by hard division method  
图 1 基于硬划分的方法划分链路质量等级

本文主要研究“好”“中等”以及“差”链路,若  $prrr$  值在  $[0, 0.2)$  内,则视为“差”链路;若  $prrr$  值在  $[0.2, 0.8)$  内,则视为“中等”链路;若  $prrr$  值在  $[0.8, 1]$  内,则视为“好”链路。

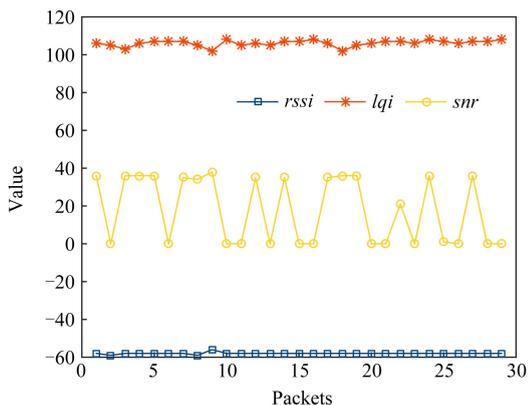
### 3 链路质量参数的确定

#### 3.1 构建链路质量参数的备选集

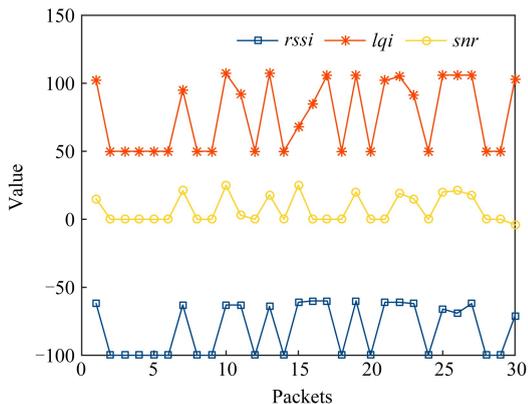
本文拟选取物理层参数为链路质量参数.基于硬件的链路质量参数,主要有 RSSI、LQI、SNR、噪音等级 (noise level, NL)、信道状态信息 (channel state information, CSI)、信号质量 (signal quality, SQ) 以及码片相关度 (chip correlation indicator, CCI) 等,其中 SQ 仅适用于直接序列展频技术的物理层调制方案<sup>[20]</sup>;CSI 适用于正交频分复用以及多人多出等技术的物理层调制方案<sup>[21]</sup>;而 RSSI 适用于所有可用的 IEEE 802.11 调制方案中.其他指标,例如 LQI,该参数通常与 RSSI 一起评估或预测链路质量<sup>[20]</sup>;通常 NL 不单独作为链路质量参数,而是与 RSSI 计算得到 SNR;LQI 被定义为符号相关度<sup>[22]</sup>即 CCI,文献[23]指出 LQI, RSSI, SNR 均与 PRR 存在着相关性.考虑到指标的通用性以及其与分类目标的相关性,本文选取 LQI, RSSI, SNR 构建链路质量参数的备选集。

从“好”“中等”以及“差”不同链路中随机选取 1 个周期并观察  $rssi, snr, lqi$  的变化,如图 2 所示.其

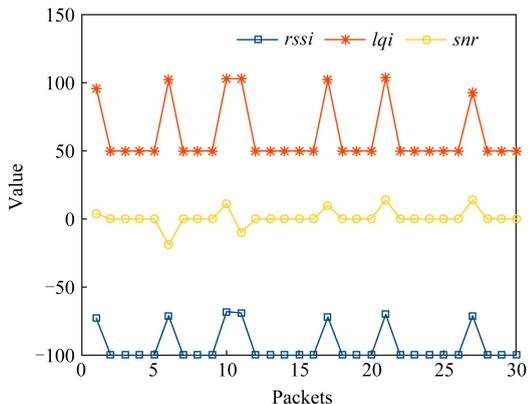
中,1 个探测周期内发送 30 个数据包,对于丢失的数据包,本文采用最小值填充;采用 0 对  $snr$  的缺失值进行填充,采用 -100 对  $rssi$  的缺失值进行填充,采用 50 对  $lqi$  的缺失值进行填充。



(a) Good link



(b) Intermediate link



(c) Bad link

Fig. 2 The value of hardware parameters in different links

图 2 周期内不同链路的硬件参数值

由图 2(a)可知,“好”链路一般位于连通区,链路质量较好,具有较高的  $lqi, rssi, snr$  值,同时  $lqi$  和  $rssi$  处于稳定状态,相比之下  $snr$  具有一定程度

的波动;而“中等”链路一般位于过渡区,链路质量一般,如图 2(b)所示, $lqi, rssi, snr$  值均有大幅度的波动,相比于连通区的链路,具有较低的  $lqi, rssi, snr$  值;“差”链路一般位于非连通区的链路,链路质量差,如图 2(c)所示, $lqi, rssi, snr$  处于稳定状态但是数值几乎为最小值.由上分析可得出结论:

若  $lqi, rssi, snr$  值越高,波动程度越低,则链路质量越好.

本文使用  $lqi, rssi, snr$  的均值  $\mu_{lqi}, \mu_{rssi}, \mu_{snr}$  表示探测周期内各硬件参数的数值大小;使用  $lqi, rssi, snr$  的极值  $r_{lqi}, r_{rssi}, r_{snr}$  与方差  $\sigma_{lqi}^2, \sigma_{rssi}^2, \sigma_{snr}^2$  表示探测周期内各硬件参数的波动程度;将  $lqi, rssi, snr$  的均值、极值与方差作为链路质量参数  $M$  的备选集(candidate set, CS),记为  $M_{CS}$ ,其中  $\mu = \{\mu_{lqi}, \mu_{rssi}, \mu_{snr}\}, r = \{r_{lqi}, r_{rssi}, r_{snr}\}, \sigma^2 = \{\sigma_{lqi}^2, \sigma_{rssi}^2, \sigma_{snr}^2\}, M_{CS} = \{\mu, r, \sigma^2\}$ .

### 3.2 链路质量参数的选取

链路质量参数的选取旨在不错过任何一个有用的链路质量参数的同时又保证链路质量参数之间不存在冗余现象,减少链路质量评估模型不必要的时间与空间的复杂度.包裹式方法将特征选择过程与训练过程整合在一起,以模型的预测或分类能力作为衡量特征子集的选择标准,从模型性能的角度出发,包裹式方法具有更优的性能,但是需花费一定的时间开销<sup>[24]</sup>,而 AGRNN 具有快速学习的能力.因此,本文采用包裹式方法,结合 AGRNN 评价各备选子集的重要性,选取链路质量参数,具体的实现过程如算法 1 所示:

**算法 1.** 基于包裹式的链路质量参数选取算法.

输入:实验训练集样本  $D$ ,链路质量参数的备选集  $M_{CS} = \{\mu, r, \sigma^2\}$ ,学习算法  $\zeta$  为 AGRNN,停止条件控制参数  $T$ ;

输出:链路质量参数的选取结果  $M$ .

- ① 初始化最小损失误差  $E_{\min} = \infty, T = |M_{CS}|$ ,  
 $M = \emptyset, t = 1$ ;
- ② for all  $m_{u,v} \in M_{CS}, u \neq v$  do
- ③ 计算  $m_u$  与  $m_v$  的相关系数  $Corr_{uv}$ ;
- ④ if  $Corr_{uv} > 0.8$
- ⑤ deleted  $m_u$ , update  $M_{CS}$ ;
- ⑥ end if
- ⑦ end for
- ⑧ while  $t \leq T$  do
- ⑨  $A = M_{CS} - M$ ;
- ⑩ for all  $m_p \in A$  do

- ⑪  $M' = M \cup m_p$ ;
- ⑫  $E = CrossValidation(\zeta(D^{M'}))$ ;
- ⑬ if ( $E < E_{\min}$ ) then
- ⑭  $E_{\min} = E$ ;
- ⑮  $M = M'$ ;
- ⑯ end if
- ⑰ end for
- ⑱  $t = t + 1$ ;
- ⑲ end while
- ⑳ return  $M$ .

AGRNN 分类性能指标  $E$  为 Loss 损失误差.序列前向选择策略只能加入特征而不能剔除特征,即便后面加入的特征与已选入的特征存在冗余,基于此本文在进行链路质量参数选取之前利用皮尔逊相关系数去除冗余参数,如步骤②~⑦,其中判断参数之间是否冗余的阈值设定为经验值 0.8,而后采用序列前向选择策略选取链路质量参数,如步骤⑧~⑳.

## 4 链路质量评估

本文采用自适应广义回归神经网络构建链路质量评估模型,AGRNN 为每个链路质量参数分配不同的光滑因子,借鉴误差反向传播的思想对光滑因子进行修正,能够快速自适应 WSNs 中动态链路环境,提高链路质量评估准确性的同时具有较快的学习速度.

### 4.1 模型结构的确定

AGRNN 是基于广义回归的神经网络,故其结构与广义回归神经网络相似.AGRNN 链路质量评估模型由 4 个部分组成如图 3 所示,分别为输入层、模式层、求和层和输出层<sup>[14]</sup>.输入层神经元个数与  $M$  中链路质量参数的个数相对应,记为  $d$ ,模式层神经元个数小于或等于训练集样本数,记为  $n$ ,本文随机将实验样本按 7:3 的比例划分为训练集样本  $D = \{(x_k, y_k)\}_{k=1}^N$  和测试集样本  $V = \{(x'_k, y'_k)\}_{k'=1}^{N'}$ ,其中  $x_k, x'_k \in \mathbb{R}^d, y_k, y'_k \in \{1, 2, 3\}$ .求和层由 2 种类型的神经元组成:一种为所有模式层神经元输出的算术之和,记为  $S_D$ ;另一种为所有模式层神经元加权求和,记为  $S_N$ <sup>[14]</sup>,其个数等于输出层神经元数.输出层神经元个数为链路质量类别数,本文研究“好”“中等”以及“差”链路,故输出层神经元个数为 3.

图 3 为 AGRNN 链路质量评估模型结构图,  $x_k = (x_k^1, \dots, x_k^d, \dots, x_k^d)$  对应评估模型的输入层;  $y_k = (y_k^1, y_k^2, y_k^3)$  为输出的编码形式,对应评估模型

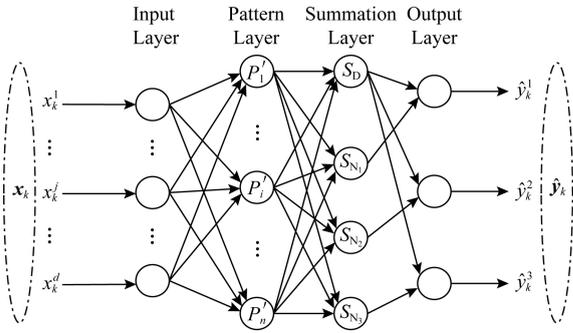


Fig. 3 The structure of LQE model based on AGRNN  
图3 基于 AGRNN 链路质量评估模型结构

的输出层,其中  $\hat{y}_k = (1, 0, 0)$ ,  $\hat{y}_k = (0, 1, 0)$ ,  $\hat{y}_k = (0, 0, 1)$  分别对应“好”“中等”以及“差”链路;模式层神经元传递函数为

$$P'_{i,q} = P'(\mathbf{x}_q, \hat{\mathbf{x}}_i) = \exp\left[-\sum_{j=1}^d \frac{(x_q^j - \hat{x}_i^j)^2}{(o^j)^2}\right], \quad (2)$$

其中,  $\{(\mathbf{x}_q, \mathbf{y}_q)\}_{q=1}^{r'}$  为模型输入,  $r' \leq N$ , 第  $i$  个模式层神经元对应样本  $(\hat{\mathbf{x}}_i, \hat{\mathbf{y}}_i)$ ,  $i \in \{1, 2, \dots, n\}$ ,  $P'_i$  为模式层中第  $i$  个神经元的输出;  $x_q^j$  为  $\mathbf{x}_q$  中第  $j$  个链路质量参数值;  $\hat{x}_i^j$  为  $\hat{\mathbf{x}}_i$  中第  $j$  个链路质量参数值;  $o^j$  为第  $j$  个链路质量参数对应的光滑因子。

链路质量评估值的输出为

$$\hat{y}_q = \frac{\sum_{i=1}^n \hat{\mathbf{y}}_i \times \exp(-P'(\mathbf{x}_q, \hat{\mathbf{x}}_i))}{\sum_{i=1}^n \exp(-P'(\mathbf{x}_q, \hat{\mathbf{x}}_i))}, \quad (3)$$

其中,  $\hat{\mathbf{y}}_i$  为第  $i$  个模式层神经元对应样本的真实链路质量等级。

本文采用的 Loss 损失函数为

$$e_q = (\hat{y}_q^1 - y_q^1)^2 + (\hat{y}_q^2 - y_q^2)^2 + (\hat{y}_q^3 - y_q^3)^2, \quad (4)$$

$\mathbf{y}_q = (y_q^1, y_q^2, y_q^3)$  为第  $q$  个被评估样本对应真实链路质量等级的编码形式。

本文则借鉴误差反向传播的思想对光滑因子进行修正为

$$o_q^j = -\eta \frac{\partial e_q}{\partial o^j} + o_{q-1}^j, \quad (5)$$

其中,  $\eta$  为学习速率因子,一般为很小的正数。

## 4.2 模型的构建

基于 AGRNN 的链路质量评估模型的训练实质就是寻找最优光滑因子  $\mathbf{o} = (o^1, \dots, o^j, \dots, o^d)$  的过程,模型的训练与评估如算法 2 所示:

**算法 2.** 基于 AGRNN 的链路质量评估模型。

输入:训练集样本  $D$ ,测试集样本  $V$ ,  $o_0$  为默认光滑因子的值(本文设为经验值 0.04),  $KFold$  为交

叉函数,  $d$  为链路质量参数的个数,学习算法  $\zeta$  为 AGRNN;

输出:测试集  $V = \{(\mathbf{x}'_{k'})\}_{k'=1}^{N'}$  的链路质量评估结果为  $\{(\hat{\mathbf{y}}'_{k'})\}_{k'=1}^{N'}$ 。

- ① 将训练集与测试集样本的输入归一化;
- ② 将训练集样本的输出转换为 one-hot 编码形式;
- ③ for all  $D_{i'} \subseteq KFold(n\_splits = 5), split(D)$  do
- ④ 将  $D_{i'}$  中的链路质量参数作为输入,余下  $D - D_{i'}$  分别为模式层神经元对应的样本,此时  $r'$  为  $D_{i'}$  的样本数,  $n$  为  $D - D_{i'}$  样本数;
- ⑤ for  $q = 1, 2, \dots, r'$  do
- ⑥ for  $i = 1, 2, \dots, n$  do
- ⑦ for  $j = 1, 2, \dots, d$  do
- ⑧ if  $q == 1$  then
- ⑨  $\hat{o}_q^j = o_0$ ;
- ⑩ else
- ⑪  $\hat{o}_q^j = \hat{o}_{q-1}^j$ ;
- ⑫ end if
- ⑬ 根据式(2)计算  $P'_{i,q}$ ;
- ⑭ end for
- ⑮ 根据式(3)计算  $\hat{\mathbf{y}}_q$ ;
- ⑯ end for
- ⑰ 根据式(4)计算  $e_q$ ;
- ⑱ 根据式(5)对光滑因子进行修正;
- ⑲ 若修正后的光滑因子不在给定的界限内,则令其为邻近的边界值;
- ⑳ end for
- ㉑ 得到局部最优光滑因子  $\hat{\mathbf{o}}_q$ ,将其加入到  $H$  中;
- ㉒ end for
- ㉓ 对光滑因子组  $H$  求均值,得到最优光滑因子  $\mathbf{o}$ ;
- ㉔ 链路质量评估结果为  $\zeta(D, V, \mathbf{o})$ ,其中  $\zeta(D, V, \mathbf{o})$  可由训练过程中式(2)、式(3)求得,此时  $V$  对应模型的输入,  $D$  中样本分别为模式层神经元对应样本,光滑因子为  $\mathbf{o}$ 。

AGRNN 为第  $j$  个链路质量参数分配光滑因子  $o^j$ ,并进行自适应修正,为了使得光滑因子不陷入局部最优,本文采用 5 折交叉验证,对 5 次训练得到的光滑因子求均值,最优光滑因子  $\mathbf{o}$  如步骤③~⑳。若没有给定光滑因子一个界限,  $o^j$  可能出现负值或者过大等情况,这样会对评估结果造成较大的影响,故使  $o^j \in [o_{\min}, o_{\max}]$  来保证评估的准确性如步骤⑲。

模型训练结束后得到最优光滑因子  $\sigma$ , 通过步骤④可得到测试集  $V = \{(\mathbf{x}'_{k'})\}_{k'=1}^{N'}$  的链路质量评估结果.

### 4.3 模型的评价

本文基于 AGRNN 构建  $M$  与链路质量等级的关系模型, 从而对链路质量进行评估, 并采用常见指标: 准确率 *Accuracy*、召回率 *Recall* 对评估模型的性能进行分类; 采用泛化误差 (generalization error, *GR*) 对评估模型的泛化能力进行评价; 采用计算时间对评估模型的时间复杂度进行评价.

准确率 *Accuracy* 为测试集样本中被正确识别的样本数与测试集样本总数之比:

$$Accuracy(f:V) = \frac{1}{N'} \sum_{k'=1}^{N'} II(f(\mathbf{x}'_{k'}) = \mathbf{y}'_{k'}), \quad (6)$$

其中,  $k'$  为第  $k'$  个测试集样本,  $k' \in \{1, 2, \dots, N'\}$ ,  $II(f(\mathbf{x}'_{k'}) = \mathbf{y}'_{k'})$  为判断测试集样本是否被正确识别的函数, 是为 1, 否为 0.

召回率 *Recall* 指的是对于某一链路质量等级测试样本, 分类正确的正样本与真正正样本数之比, 其中真正正样本数为:  $TP$  以及真实标签为该等级但未被正确识别为该等级的样本数  $FN$  之和, 通常用来衡量分类器的查全率:

$$Recall = \frac{TP}{TP + FN}, \quad (7)$$

若计算“好”链路的召回率,  $TP$  与  $FN$  之和表示测试集样本中真实的链路质量等级为“好”的样本数,  $TP$  表示测试集样本中被评估为“好”链路且被正确识别的样本数.

学习方法的泛化能力是指模型适应新样本的能力, 通常采用 *GR* 进行评价, *GR* 表示测试误差与训练误差的差异程度<sup>[25]</sup>.

文献[26]采用模型的学习时间作为评价指标, 在此基础上本文增加评估所花费的时间对模型的时间复杂度进行评价, 即计算时间 (computing time, *CT*), 该指标由评估模型训练和评估所花费的总时间进行表征.

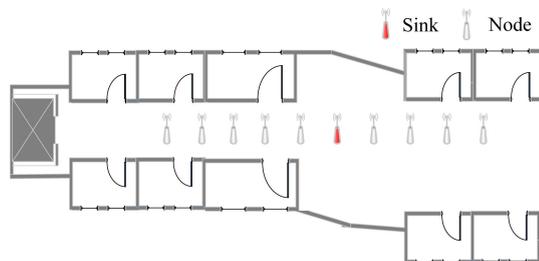
## 5 实验结果与分析

本文实验在 TelosB 节点上进行, 采用实验室自主研发的链路质量测试平台对节点间的链路质量信息进行收集、展示以及离线分析, 采用 Python 进行实验床实验, 选取基于多项式法、随机森林、支持向量分类器的链路质量评估模型与本文提出的链路质

量评估模型进行对比, 并对不同场景下各模型的评价结果进行分析.

### 5.1 实验场景

本文分别在室内、公园和公路实验场景下部署 8~9 个节点组成网络进行实验, 如图 4 所示, 位于中间的节点为 Sink 节点, Node 表示感知节点, 实验参数设置如表 1 所示.



(a) The indoor scenario



(b) The park scenario



(c) The motorway scenario

Fig. 4 The experiment scenarios

图 4 实验场景图

室内实验场景如图 4(a) 所示, Sink 节点被放置在中间, 在 Sink 节点前、后 2 个方向分别部署 4 个和 5 个等距的感知节点, 该场景设置在研究生实验室外的廊道. 公园实验场景如图 4(b) 所示, Sink 节点被放置在中间, 在 Sink 节点上、下、左、右 4 个方向分别部署 2 个感知节点, 该场景设置在校园的小树林. 公路实验场景如图 4(c) 所示, 节点摆放与公园场景类似, 该场景设置在流量较大的路段. 具体实验参数设置如表 1 所示:

**Table 1 The Setting of Experimental Parameters****表 1 实验参数设置**

Parameters Attribute	Value of Parameters
Channel	26
Number of Packets	30
Packet Sending Interval/s	0.2
Test Period/s	12
Transmission Speed/Kbps	250
Monitoring	Active

## 5.2 模型的验证与对比

为验证基于 AGRNN 的链路质量评估模型具有较优的评估性能、较好的泛化能力以及较低的时间复杂度,本文分别在室内、公园以及公路 3 个场景中收集实验样本,选取  $M = \{m_1, \dots, m_j, \dots, m_d\}$  作为评估模型的输入,真实的链路质量等级为标签,选取准确率、召回率、泛化误差、计算时间为评价指标,采用交叉验证的方法优化 SVC 参数从而构建基于 SVC 的链路质量评估模型<sup>[12]</sup>;采用网格搜索法确定变化维度从而构建基于多项式回归的链路质量评估模型<sup>[9]</sup>;采用 Bagging 集成的思想构建基于随机森林的链路质量评估模型<sup>[13]</sup>,将上述评估模型与本文提出的基于 AGRNN 的链路质量评估模型进行对比。

### 1) 准确率对比分析

基于 AGRNN 的评估值为所有训练集样本的加权平均,权重因子为训练集样本与测试集样本对应输入的欧氏距离平方指数,网络收敛于样本量积聚最多的优化回归面,AGRNN 在分类能力上具有较强优势,如图 5 所示:

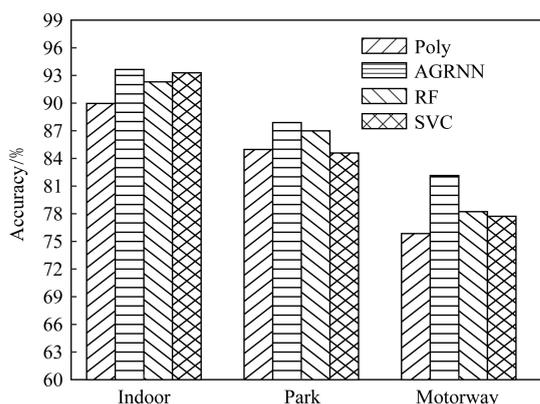


Fig. 5 The accuracy of different LQEs in different scenarios

图 5 不同场景下链路质量评估模型的准确率

图 5 为不同场景下链路质量评估模型的准确

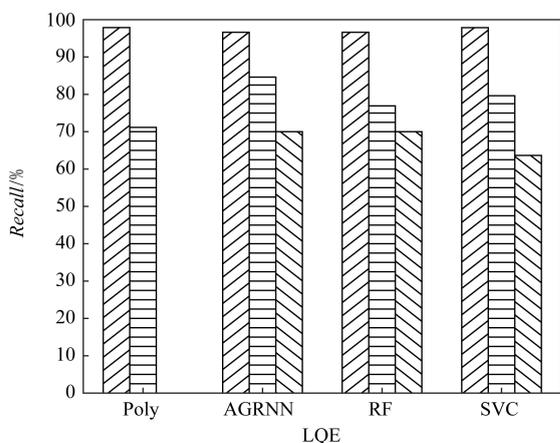
率,可知公路实验场景下的评估准确率远低于室内实验场景的准确率,可能是由于收集公路实验场景的链路信息时处于流量高峰期,存在移动障碍物影响和干扰,链路具有动态、不稳定的特点使得评估模型在训练过程中难以得到最优解,但是基于 AGRNN 的链路质量评估模型获得了最高的准确率,为 82.1%,主要是因为 AGRNN 可自适应地为每个链路质量参数分配不同的光滑因子,使该评估模型能较为准确地评估链路质量.在公园实验场景下,基于 AGRNN 的链路质量评估模型的准确率(87.8%)远高于基于 Poly 和基于 SVC 的链路质量评估模型(85.8%, 84.5%),略高于基于 RF 的链路质量评估模型的准确率(86.9%),主要是因为公园实验场景下,电脑处于无源状态,如图 4(b)所示, Sink 节点插在电脑上,故收集的实验数据量(852 条)远少于室内场景(2018 条),仅通过少量的训练集样本,单一的分类器难以准确地评估该实验场景下的链路质量,基于 RF 的链路质量评估模型是以随机树为基分类器的一种集成学习模型,评估该实验场景下的链路质量更具优势,而 AGRNN 在样本数据较少时仍具有较好的分类或预测能力<sup>[14]</sup>,故此在该实验场景下能相对较准确地评估链路质量.在室内场景下,基于 AGRNN 以及基于 SVC 的链路质量评估模型均具有较高的准确率分别为 93.6%和 93.2%。

在不同的实验场景中,基于 AGRNN 的链路质量评估模型其准确率优于基于 RF、基于 Poly、基于 SVC 的 3 个链路质量评估模型,且在样本数据较少的情况下仍能较准确地评估链路质量。

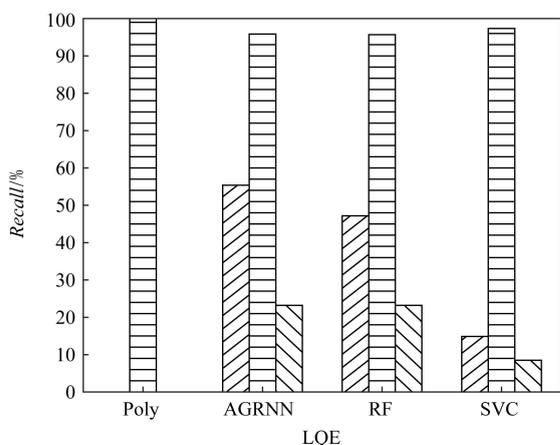
### 2) 召回率对比分析

本文采用召回率来评价各链路质量评估模型正确识别“好”“中等”以及“差”链路的能力,评价结果如图 6 所示。

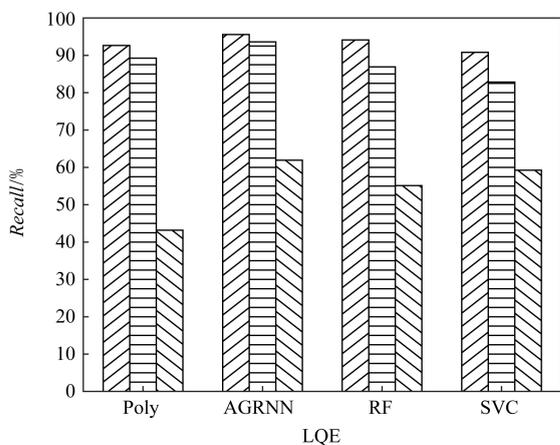
图 6 中 Level 1, Level 2, Level 3 分别表示链路质量等级 1, 2, 3, 对应“好”“中等”以及“差”链路.由图 6(a)可知,在室内实验场景中,基于 AGRNN 及基于 RF 的 2 个链路质量评估模型在等级 1 和 3 下的召回率相似且具有较高值分别为 96%和 70%,但在等级 2 下基于 AGRNN 的链路质量评估模型的召回率比基于 RF 的链路质量评估模型高了 7.7 个百分点,可知在该实验场景下基于 AGRNN 的链路质量评估模型具有更高的召回率.如图 6(b)所示,在公园实验场景下基于 Poly 的链路质量评估模型在等级 1 和 3 下的召回率为 0,可能是因为收集该场景下的实验样本时链路质量主要集中在等级 2 上,



(a) The indoor scenario



(b) The park scenario



(c) The motorway scenario

▨ Level 1    ▤ Level 2    ▧ Level 3

Fig. 6 The recall of different LQEs in different scenarios

图 6 不同实验场景下链路质量评估模型的召回率

基于 AGRNN 以及基于 RF 的 2 个链路质量评估模型在链路质量等级 3 下的召回率均为 23%，且 AGRNN 的链路质量评估模型在链路质量等级 1 下的召回率较 RF 高了 7.9 个百分点，可知基于 AGRNN

的链路质量评估模型正确识别少数类样本的能力较强。在公路实验场景下，如图 6(c)所示，基于 AGRNN 的链路质量评估模型在不同链路质量等级下的召回率均最优分别为 95%，90%，60%。

综上可知基于 AGRNN 的链路质量评估模型能够更好地识别少数类样本同时在不同链路质量等级下具有更高的查全率。

通过对各评估模型的准确率和召回率的分析可知：在不同实验场景下，基于 AGRNN 的链路质量评估模型具有较高的准确率和召回率，即具有更优的评估性能。

### 3) 泛化误差对比分析

本文采用  $GR$  来评价各评估模型的泛化能力，若  $GR$  越小说明模型评估未知数据的能力越强。本文分别计算各评估模型在不同实验场景下的  $GR$  值，如表 2 所示：

Table 2  $GR$  of Different LQEs in Different Scenarios表 2 不同实验场景下各链路质量评估模型的  $GR$  值

LQE	Indoor	Park	Motorway
Poly	0.010	0.308	0.092
AGRNN	0.130	0.078	0.172
RF	0.147	0.238	0.202
SVC	0.175	0.017	0.234

由表 2 可知，基于 Poly 的链路质量评估模型在室内、公路实验场景下的泛化能力优于其他评估模型，但是在公园实验场景下其泛化误差(0.308)为基于 SVC 的链路质量评估模型泛化误差的 18 倍；基于 SVC 的链路质量评估模型在公园场景下泛化误差为 0.017，可知在该场景下其泛化性能力相比于其他评估模型具有较大的优势，但是在室内和公路实验场景下，其泛化误差分别为基于 Poly 的链路质量评估模型泛化误差的 17.5 倍和 2.5 倍，并没有体现出其较好的泛化能力；基于 AGRNN 的链路质量评估模型在室内和公路实验场景下的泛化误差分别为基于 Poly 的链路质量评估模型泛化误差的 13 倍和 1.87 倍，且在公园实验场景下基于 AGRNN 的链路质量评估模型的泛化误差仅为基于 SVC 的链路质量评估模型泛化误差的 4.6 倍。

综合 3 个实验场景下各评估模型的泛化误差可知：基于 AGRNN 的链路质量评估模型的泛化误差一直较为稳定且接近最小值，具有较好的泛化能力。

### 4) 计算速度对比分析

本文采用计算时间评价各链路质量评估模型的

时间复杂度.实验样本从 1 000 条增加到 10 000 条,其中步长设为 1 000,作为各链路质量评估模型的输入.在训练基于 RF 的链路质量评估模型时决策树个数为 500.采用 5 折交叉验证对基于 AGRNN 以及 SVC 链路质量评估模型进行训练,并对测试集样本进行评估,最后得到计算时间,如图 7 所示:

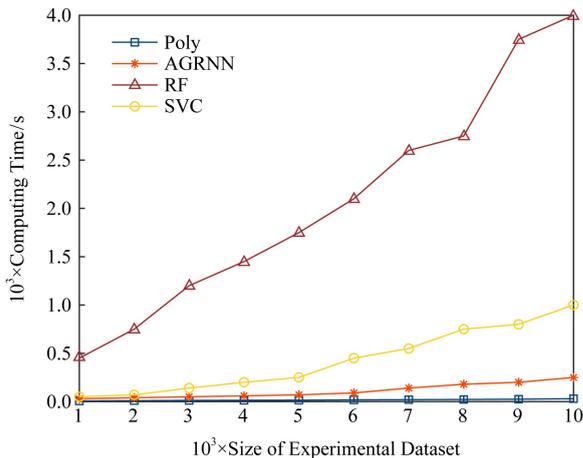


Fig. 7 Computing time of different LQEs

图 7 各链路质量评估模型的计算时间图

由图 7 可知,计算时间随着实验数据量的增加而增加.基于 RF 的链路质量评估模型,其计算时间呈线性增加,且其计算时间远远大于其他评估模型,可能是因为该方法是基于 Boosting 集成的思想,而其他评估模型都是单分类器;在实验样本大于 5 000 时基于 SVC 的链路质量评估模型其计算时间显著增加,而基于 AGRNN 的链路质量评估模型,其计算时间均在 250 s 以内,略高于基于 Poly 的链路质量评估模型,但仍处于一个较低的水平,该结果验证了自适应广义回归神经网络具有快速学习的能力——它与链路质量评估相结合可以减少模型的计算时间,降低时间复杂度.

## 6 总 结

如何选择合适的链路质量参数,构建具有较优评估性能、较好泛化能力和较低时间复杂度的链路质量评估模型,仍是一个具有挑战的任务.本文提出了基于包裹式的链路质量参数选取算法,采用 AGRNN 评价各备选子集的重要性,选取链路质量参数;在链路质量的评估问题中,人工神经网络取得显著效果且基于映射关系的链路质量评估方法综合考虑链路质量评估的特性,本文采用 AGRNN 构建

链路质量参数与链路质量等级的关系模型,提出基于 AGRNN 的链路质量评估模型.与基于 SVC、基于 Poly、基于 RF 的 3 个链路质量评估模型对比可知:在 3 个不同的实验场景下,本文提出的评估模型其准确率高于其他评估模型;对于少数类样本来说,本文提出的模型能更全面地找出该类样本同时在不同链路质量等级下具有更高的召回率.综合 3 个实验场景下各评估模型的泛化误差可知:基于 AGRNN 的链路质量评估模型具有较好的泛化性;随着样本数的增加,本文提出的模型其计算时间虽略高于基于 Poly 的链路质量评估模型,但远远低于其他 2 个链路质量评估模型.基于 AGRNN 的链路质量评估模型具有较快的学习速度以及较强自适应能力,未来可将该模型运用于在线的链路质量评估.

## 参 考 文 献

- [1] Xie Chenghan, Lu Saijie, Wang Hao, et al. Output feedback control based on event-based sample in wireless sensor networks [J]. *Journal of Computer Research and Development*, 2017, 54(11): 2639-2645 (in Chinese) (谢承翰, 陆赛杰, 王皓, 等. 无线传感器网络中基于事件驱动的输出反馈控制[J]. *计算机研究与发展*, 2017, 54(11): 2639-2645)
- [2] Baccour N, Koubaa A, Mottola L, et al. Radio link quality estimation in wireless sensor networks: A survey [J]. *ACM Transactions on Sensor Networks*, 2012, 8(4): 192-224
- [3] Weng Lina, Liu Yiming, Liu Lei, et al. Wireless link quality estimation and prediction: A survey [J]. *Journal of China Academy of Electronics and Information Technology*, 2016, 11(3): 239-267 (in Chinese) (翁丽娜, 刘轶铭, 刘磊, 等. 无线链路质量评估及预测方法综述[J]. *中国电子科学研究院学报*, 2016, 11(3): 239-267)
- [4] Audéoud H J, Heusse M. Quick and efficient link quality estimation in wireless sensors networks [C] //Proc of the 14th Annual Conf on Wireless On-demand Network Systems and Services(WONS'18). Piscataway, NJ: IEEE, 2018: 87-90
- [5] Qin Fei, Zhang Qilong, Zhang Wuxiong, et al. Link quality estimation in industrial temporal fading channel with augmented Kalman filter [J]. *IEEE Transactions on Industrial Informatics*, 2019, 15(4): 1936-1946
- [6] Baccour N, Koubaa A, Youssef H, et al. F-LQE: A fuzzy link quality estimator for wireless sensor networks [C] //Proc of the 7th European Conf on Wireless Sensor Networks (EWSN'10). New York: ACM, 2010: 240-255
- [7] Deb M, Roy S, Saha B, et al. Designing a new link quality estimator for sensor nodes by combining available estimators [C] //Proc of the 7th Int Advanced Computing Conf (IACC'17). Piscataway, NJ: IEEE, 2017: 179-183

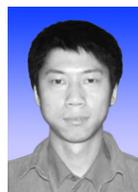
- [8] Aswale S, Ghorpade V R. Geographic multipath routing based on triangle link quality metric with minimum inter-path interference for wireless multimedia sensor networks[J/OL]. Journal of King Saud University-Computer and Information Sciences, 2018 [2019-12-09]. <https://doi.org/10.1016/j.jksuci.2018.02.001>
- [9] Gomes R D, Queiroz D V, Lima Filho A C, et al. Real-time link quality estimation for industrial wireless sensor networks using dedicated nodes [J]. Ad Hoc Networks, 2017, 59: 116-133
- [10] Kulin M, Poorter E, Kazaz T, et al. Poster: Towards a cognitive MAC layer: Predicting the MAC-level performance in dynamic WSN using machine learning [C] //Proc of the 14th Int Conf on Embedded Wireless Systems and Networks (EWSN'17). New York: ACM, 2017: 214-215
- [11] Sun Wei, Lu Wei, Li Qiyue, et al. WNN-LQE: Wavelet-neural-network-based link quality estimation for smart grid WSNs [J]. IEEE Access, 2017, 5: 12788-12797
- [12] Shu Jian, Liu Song, Liu Linlan, et al. Research on link quality estimation mechanism for wireless sensor networks based on support vector machine [J]. Chinese Journal of Electronics, 2017, 26(2): 377-384
- [13] Cerar G, Mohorcic M, Gale T, et al. Analysis of machine learning for link quality estimation [J]. arXiv preprint, arXiv:1812.08856, 2019
- [14] Specht D F. A general regression neural network [J]. IEEE Transactions on Neural Networks, 1991, 2(6): 568-576
- [15] Jondhale S R, Deshpande R S. Kalman filtering framework-based real time target tracking in wireless sensor networks using generalized regression neural networks [J]. IEEE Sensors Journal, 2019, 19(1): 224-233
- [16] Sharma S, Mukherjee A, Dongre A, et al. Ultra low power sensor node for security applications, facilitated by algorithm-architecture co-design [C] //Proc of the 30th IEEE Int Conf on VLSI Design and the 16th Int Conf on Embedded Systems (VLSID'17). Piscataway, NJ: IEEE, 2017: 101-106
- [17] Seng T L, Khalid M, Yusof R, et al. Adaptive GRNN for the modelling of dynamic plants [C] //Proc of the 17th IEEE Inter Symp on Intelligent Control (ISIC'02). Piscataway, NJ: IEEE, 2002: 217-222
- [18] Bildea A, Alphanand O, Rousseau F, et al. Link quality estimation with Gilbert-Elliot model for wireless sensor networks [C] //Proc of the 26th IEEE Inter Symp on Personal, Indoor, and Mobile Radio Communications (PIMRC'15). Piscataway, NJ: IEEE, 2015: 2049-2054
- [19] Luo Xionghui, Liu Linlan, Shu Jian, et al. Link quality estimation method for wireless sensor networks based on stacked autoencoder [J]. IEEE Access, 2019, 7(1): 21572-21583
- [20] Lowrance C J, Lauf A P. Link quality estimation in ad hoc and mesh networks: A survey and future directions [J]. Wireless Personal Communications, 2017, 96(1): 475-508
- [21] Heinzer P, Lenders V, Legendre F. Fast and accurate packet delivery estimation based on DSSS chip errors [C] //Proc of the 31st IEEE Int Conf on Computer Communications (INFOCOM'12). Piscataway, NJ: IEEE, 2012: 2916-2920
- [22] Texas Instruments. 2.4 GHz IEEE 802.15.4/ZigBee ready RF transceiver [EB/OL]. 2007 [2019-08-22]. <http://www.ti.com/cn/lit/ds/symlink/cc2420.pdf>
- [23] Liu Linlan, Gao Shengrong, Shu Jian. Link quality prediction based on random forest [J]. Journal on Communications, 2019, 40(4): 202-211 (in Chinese)  
(刘琳岚, 高声荣, 舒坚. 基于随机森林的链路质量预测[J]. 通信学报, 2019, 40(4): 202-211)
- [24] Agor J, Özalın O Y. Feature selection for classification models via bilevel optimization [J]. Computers & Operations Research, 2019, 106: 156-168
- [25] Zhang Chiyuan, Bengio S, Hardt M. Understanding deep learning requires rethinking generalization [C/OL] //Proc of the 5th Int Conf on Learning Representation (ICLR'17). 2017 [2019-12-09]. <https://arxiv.org/abs/1611.03530>
- [26] Ai Ke, Ma Guoshuai, Yang Kaikai et al. A classification method of scientific collaborator potential prediction based on ensemble learning [J]. Journal of Computer Research and Development, 2019, 56(7): 1383-1395 (in Chinese)  
(艾科, 马国帅, 杨凯凯, 等. 一种基于集成学习的科研合作者潜力预测分类方法[J]. 计算机研究与发展, 2019, 56(7): 1383-1395)



**Shu Jian**, born in 1964. Received his MSc degree in computer networks from North-western Polytechnical University. Professor in Nanchang Hangkong University, China. Senior member of CCF. His main research interests include wireless sensor networks, embedded systems, and software engineering.



**Gao Su**, born in 1994. Master candidate in Nanchang Hangkong University. Student member of CCF. Her main research interests include wireless sensor networks. (1716085212004@stu.nchu.edu.cn)



**Chen Yubin**, born in 1977. Received his MSc degree in computer software and theory from Nanchang University. Lecturer in the School of Software, Nanchang Hangkong University. His main research interests include embedded system and wireless sensor network. (chenyubin@nchu.edu.cn)