

• 物联网 •

DOI:10.15961/j.jsuese.202100951



基于门控循环单元的链路质量预测

刘琳岚¹, 肖庭忠¹, 舒 坚^{2*}, 牛明晓¹

(1.南昌航空大学 信息工程学院, 江西 南昌 330063; 2.南昌航空大学 软件学院, 江西 南昌 330063)

摘 要: 无线传感器网络中, 节点传输数据时容易受到环境中噪声的干扰, 使传输链路质量变差, 导致数据包丢失、消息重发, 从而加速节点能量的消耗, 缩短网络寿命。链路质量预测可以为上层路由协议选择高质量的无线链路进行通信提供依据, 通过链路质量预测选取高质量的链路传输数据, 可以提高数据传输效率, 减少重传次数。本文提出基于门控循环单元(gate recurrent unit, GRU)的链路质量预测方法。首先, 采用Gap Statistic算法优化的K-means++算法, 依据包接收率的分布自适应划分链路质量等级, 获得链路质量样本标签; 其次, 选择接收信号强度均值、链路质量指示均值及信噪比均值作为输入, 借助CatBoost在分类问题上的优势, 构建链路质量评估模型, 并采用网格搜索法对参数寻优; 最后, 使用滑动时间窗口构建链路质量等级时序样本集, 使用GRU提取链路质量等级时间序列的时序信息, 为进一步提高预测的准确率, 采用支持向量回归机构建链路质量预测模型, 预测下一时刻链路质量等级。本文采用真实场景中的数据进行实验, 根据主要干扰源不同, 选择实验室、走廊和停车场3个场景收集数据, 使用均方误差评价链路质量预测模型的有效性。实验结果表明, 与小波神经网络、循环神经网络和随机向量函数链等方法相比, 所提方法具有更小的预测误差, 可以准确预测链路质量等级。

关键词: 无线传感器网络; 链路质量预测; 门控循环单元; Gap Statistic算法; 链路质量等级

中图分类号: TP391

文献标志码: A

文章编号: 2096-3246(2022)06-0051-08

Link Quality Prediction Based on Gate Recurrent Unit

LIU Linlan¹, XIAO Tingzhong¹, SHU Jian^{2*}, NIU Mingxiao¹

(1.School of Info. Eng., Nanchang Hangkong Univ., Nanchang 330063, China;

2.School of Software, Nanchang Hangkong Univ., Nanchang 330063, China)

Abstract: In wireless sensor networks, nodes are susceptible to noise interference, which makes the transmission link quality worse, leading to packet loss and retransmission, thus speeding up the energy consumption of nodes and reducing the network life. Link quality prediction can provide a basis for upper layer routing protocols to select high-quality wireless links for communication, which can improve data transmission efficiency and reduce the number of retransmissions. A link quality prediction method based on gate current unit (GRU) was proposed in this paper. Specifically, the K-means++ algorithm optimized by the gap statistic algorithm was used to adaptively divide link quality levels according to the packet reception rate. In this way, the labels of link quality samples were obtained. Then, the received signal strength indicator mean, the link quality indicator mean and the signal to noise ratio mean were selected as the input. A link quality estimation model was constructed based on CatBoost because of its strength in classification, and the grid search optimization algorithm was employed to optimize the parameters of estimation model. Finally, the link quality level time series sample was constructed by sliding windows according to the results of the estimation model. GRU was used to extract link quality time series information. In order to improve the accuracy of prediction, the timing information was input to

收稿日期:2021-09-20

基金项目:国家自然科学基金项目(61962037; 62062050); 江西省自然科学基金重点项目(20202BABL202039); 江西省研究生创新专项项目(YC2020-S543)

作者简介:刘琳岚(1964—), 女, 教授. 研究方向: 物联网技术; 无线传感器网络等. E-mail: liulinlan@nchu.edu.cn

* 通信作者:舒 坚, 教授, E-mail: shujian@nch.edu.cn

网络出版时间:2022-07-26 11:32:54

网络出版地址:https://kns.cnki.net/kcms/detail/51.1773.tb.20220725.0955.014.html

support vector regression to predict the link quality level at the next moment. Data collected from the real world were used in this paper. According to the different common interference types, three scenes of lab, corridor and parking lot were taken as experiments scenarios. The Mean square error was used to evaluate the performance of link quality prediction models, and the experimental results showed that compared with wavelet neural network, recurrent neural network, and random vector functional link and so on, the proposed method can predict the link quality level accurately.

Key words: wireless sensor network; link quality prediction; gated recurrent unit; Gap Statistic algorithm; link quality level

无线传感器网络(wireless sensor network, WSN)是一种由大量具有感知能力和通信能力的微型节点部署在感知区域形成的网络^[1],在环境监测^[2]、智能交通^[3]、智能家居^[4]等领域广泛应用。通过预测链路质量,选择高质量链路进行数据传输,可以减少数据丢失和数据重传次数,提高网络的吞吐量^[5-6],因此,链路质量预测具有重要意义。

链路质量预测的方法主要包括:基于链路特性的方法、基于概率估计的方法和基于机器学习的方法^[7]。基于链路特性的方法主要采用接收信号强度指标(received signal strength indication, RSSI)、链路质量指标(link quality indicator, LQI)或信噪比(signal noise ratio, SNR)等物理层参数预测链路质量,这些参数可以直接获取,不需额外的计算,能耗低,但这些参数均依赖于节点,而硬件本身存在校准误差,且忽略了分组丢失,易产生高估链路质量的问题。基于概率估计的方法通过发送大量探测包,接收端计算包接收率(packet reception rate, PRR)等信息来预测链路质量^[8],但是该方法需要发送大量探测包,不仅增加了额外的通信开销,且对链路状况的反应不够灵敏。

基于机器学习的链路质量预测方法主要通过将链路质量预测问题转化为时序预测问题,利用机器学习或深度学习的方法挖掘训练样本的特征,深入学习模型输入与输出之间的潜在关系,利用得到的模型对下一时刻的链路质量进行预测。Liu等^[9]提出4C(Foresee),将LQI和PRR输入逻辑回归模型,预测下一时刻的PRR,然而该方法的预测准确率并不高;Liu等^[10]提出一种预测大波动链路质量的方法,命名为WNN_LQP(wavelet neural network based link quality prediction),该方法利用过渡区链路质量指标的高分辨率以及小波神经网络的学习能力预测下一时刻的链路质量,然而对于相对较小的波动,该模型无法进行准确预测;Xue等^[11]提出一种基于随机向量函数链的链路质量预测方法(link quality prediction based on random vector functional link, RVFL_LQP),选择SNR为链路质量参数,利用随机向量函数链建立时变序列和随机序列方差,对SNR的概率保证区间边界进行预测,但在真实应用中该结果无法快速反映

链路是否可用;Xu等^[12]提出一种基于循环神经网络(recurrent neural network, RNN)的链路质量预测方法(recurrent neural network based link quality prediction, RNN_LQP),降低了预测误差,但RNN内部结构复杂、训练周期长,且易出现梯度消失、梯度爆炸等问题^[13],导致RNN_LQP模型复杂度高,模型效果不稳定。采用机器学习的方法能够有效地挖掘出训练样本的特征,相比基于链路特性的方法和基于概率估计的方法,具有更高的准确性和稳定性,是目前WSN链路质量预测的主流方法,因此本文使用机器学习的方法进行链路质量预测研究。

GRU作为RNN的一种变体,在时序问题建模方面比其他机器学习方法更有优势^[14],且已被证明在时序预测问题上具有更好的预测效果^[15]。现有研究中,Abdel-Nasser等^[16]将GRU应用于无线社交网络中的链路质量预测中,并取得不错的效果。

因此,本文提出一种基于门控循环单元(gate recurrent unit, GRU)的链路质量预测方法,具体可分为3个步骤:首先,采用Gap Statistic算法优化的K-means++聚类算法,自适应地划分链路质量等级,避免由于人为选定临界点带来的误差干扰,获得更加贴近真实的链路质量样本标签。其次,采用CatBoost构建链路质量评估模型,对当前时刻的链路质量进行评估。最后,采用GRU构建链路质量预测模型,将评估模型的结果作为输入,借助GRU在时序问题上的优势提取链路质量等级的时序信息,采用支持向量回归机(support vector regression, SVR)预测得到下一时刻链路质量等级,提高链路质量预测的准确率。

1 链路质量等级划分

本文采用Gap Statistic算法^[17]优化的K-means++算法^[18]对PRR进行聚类,划分链路质量等级,获得链路质量样本标签。

1.1 Gap Statistic算法

Gap Statistic算法以样本集的统计规律为基础,通过间隙统计量刻画样本观察值和参考分布下期望值之间的差异,得到最佳聚类数。本文采用Gap Statistic算法确定聚类数目,具体过程描述如下:

1)对于任意两个PRR样本 p_m 和 p_m ,计算它们之

间的距离 $d_{mm'}$:

$$d_{mm'} = \|p_m - p_{m'}\|^2 \quad (1)$$

式中: $m = 1, 2, \dots, N$; $m' = 1, 2, \dots, N$; $m \neq m'$; N 为PRR样本数。

2) 假设将PRR样本集分为 K 类 $\{C_1, C_2, \dots, C_k, \dots, C_K\}$, 计算第 k 类中样本距离总和 D_k :

$$D_k = \sum_{m=1}^{|C_k|} \sum_{m'=1}^{|C_k|} d_{mm'}, m \neq m' \quad (2)$$

式中, $|C_k|$ 为第 k 类样本数目。

计算所有类的样本距离的平均和 W_K :

$$W_K = \sum_{k=1}^K \frac{1}{2|C_k|} D_k \quad (3)$$

3) 对 W_K 取对数, 并与使用蒙特卡洛模拟对数的期望进行比较, 如式(4)所示:

$$Gap(K) = E^*(\text{lb}(W_K)) - \text{lb}(W_K) \quad (4)$$

式中, $E^*(\text{lb}(W_K))$ 为使用蒙特卡洛模拟得到的 $\text{lb}(W_K)$ 的期望。

4) 取使式(5)成立的最小值 K 为PRR样本集的最佳聚类数目。

$$Gap(K) - Gap(K+1) + s_{K+1} \geq 0 \quad (5)$$

式中, s_{K+1} 为修正后的标准差, 其表达式为:

$$s_{K+1} = s'_{K+1} \sqrt{1 + 1/B} \quad (6)$$

式中, s'_{K+1} 为使用蒙特卡洛模拟得到的 $\text{lb}(W_{K+1})$ 的标准差, B 为蒙特卡洛模拟中生成参考数据集的数量。

1.2 改进的K-means++算法

K-means算法实现简单, 聚类效果好, 非常容易部署实施, 在机器视觉、地质统计学等领域得到成功的应用^[19]。K-means算法需要提前给出聚类数目 K , 但在不同的数据集中 K 的取值很难确定^[20], 由于随机选择 K 个样本点作为聚类中心, 不同的初始聚类中心可能导致不同的聚类结果^[21]。与K-means算法不同, K-means++算法在选取聚类中心时, 尽量使聚类中心互相离得远, 可减小聚类结果的波动。

本文采用Gap Statistic算法确定聚类数目 K , 再采用K-means++算法划分链路质量等级, 具体过程如下:

1) 采用Gap Statistic算法确定 K 。

2) 从PRR样本集中随机选择一个样本点作为第一个初始化聚类中心。

3) 计算PRR样本与它最近聚类中心的欧式距离 $D(p_m)$:

$$D(p_m) = \sqrt{(p_m - p)^2} \quad (7)$$

式中, p 为距离 p_m 最近的聚类中心。

4) 计算样本被选为下一个聚类中心的概率 $Q(p_m)$, 如式(8)所示, 并使用轮盘法选出下一个聚类中心。

$$Q(p_m) = \frac{D(p_m)^2}{\sum_{m=1}^N D(p_m)^2} \quad (8)$$

5) 重复步骤3)和4), 直到选出 K 个聚类中心。

6) 将每个样本点归类到离它最近聚类中心所属的类。

7) 重新计算每个聚类内PRR的均值, 将其定为新的聚类中心。

8) 重复步骤6)和7), 直至聚类中心不再变化。

9) 得到每个样本 x_i 的链路质量等级 y_i 。

2 链路质量评估

以链路质量参数作为输入, 第1节得到的链路质量等级作为输出, 采用CatBoost构建链路质量评估模型, 主要包括评估样本的构建、样本预处理及评估模型的构建。

2.1 评估样本的构建

本文选取RSSI均值、LQI均值和SNR均值为链路质量参数, 采用链路质量等级评价链路质量。链路质量样本集为:

$$\{(x_1, y_1), (x_2, y_2), \dots, (x_i, y_i), \dots, (x_N, y_N)\} \quad (9)$$

式中: $x_i = (\bar{R}_i, \bar{L}_i, \bar{S}_i)$, 其中, $\bar{R}_i$ 为样本 x_i 的RSSI均值, $\bar{L}_i$ 为样本 x_i 的LQI均值, $\bar{S}_i$ 为样本 x_i 的SNR均值; y_i 为样本 x_i 的链路质量等级; N 为样本数。

2.2 样本的预处理

由于传感器节点在收发信息时会受到多径效应、信道、温湿度或障碍物等的干扰, 收集到的数据存在缺失值, 本文计算缺失值样本在一个发送周期内的均值, 采用均值填充法对样本进行预处理。

2.3 基于CatBoost的链路质量评估模型的构建

CatBoost是由Yandex的研究人员和工程师开发的基于梯度提升决策树(gradient boosting decision tree, GBDT)的机器学习模型^[22], 采用排序提升算法替换GBDT中的梯度估计方法^[23], 能够减少GBDT算法中因梯度偏差和预测偏移而导致的过拟合问题, 有着更好的泛化能力^[24]。

本文提出基于CatBoost的链路质量评估模型(link quality estimation based on CatBoost, CatBoost_LQE), 其中, 排序提升算法描述如下:

1) 对训练集中 τ 个样本对进行随机排列, 将序列标记为 ψ 。

2) 对于每一个样本对 (x_i, y_i) , 初始化该样本对的模型 M_i 。

3) 对于每一轮迭代:

① 计算样本 x_i 的无偏度梯度估计:

$$g_i = y_i - M_{\psi(i)-1}(x_i) \quad (10)$$

② 对于每一个样本对 (x_i, y_i) , 使用该样本对在序列 ψ 之前的训练集 $(x_j, g_j) : \psi(j) \leq i$ 训练下一个分类回归树 ΔM , 将其更新到模型中, 如式 (11) 所示:

$$M_i \leftarrow M_i + \Delta M \quad (11)$$

4) 迭代完成, 输出 M_r 即为最终模型。

CatBoost 通过一组弱学习器迭代得到一个强学习器, 每一轮迭代生成一棵决策树。本文输入链路质量样本对 (x_i, y_i) , 通过 e 轮迭代构建链路质量评估模型。

设第 $e-1$ 轮迭代得到的学习器为 $F^{e-1}(\cdot)$, 计算 $F^{e-1}(\cdot)$ 的梯度 $g^{e-1}(x_i, y_i)$:

$$g^{e-1}(x_i, y_i) = \frac{\partial E(y_i, F^{e-1}(x_i))}{\partial F^{e-1}(x_i)} \quad (12)$$

式中, $E(y_i, F^{e-1}(x_i))$ 为 $F^{e-1}(\cdot)$ 的损失。

将第 $e-1$ 轮的负梯度作为残差, 传递到第 e 轮, 找到一个新的弱学习器 $h^e(\cdot)$, 使损失值最小, $h^e(\cdot)$ 表达式如下:

$$h^e(\cdot) = \arg \min E(y_i, F^{e-1}(x_i) + h^e(x_i)) \quad (13)$$

得到最终的强学习器 $F^e(\cdot)$:

$$F^e(\cdot) = F^{e-1}(\cdot) + h^e(\cdot) \quad (14)$$

3 链路质量预测

根据链路质量在短期内表现出强时间相关性的特点, 本文采用 CatBoost_LQE 得到的链路质量等级时间序列作为输入, 采用滑动窗口法确定输入序列的长度, 进而构建基于 GRU 的链路质量预测模型。

3.1 预测样本的构建

采用 CatBoost_LQE 对历史链路质量样本进行评估, 得到链路质量等级值序列 $\{l_1, l_2, \dots, l_i, \dots, l_n\}$, 其中, l_i 为第 i 个时刻链路质量等级值, n 为样本数。使用大小为 w 的时间窗口构建时序样本集:

$$\{(l_1, \phi_1), (l_2, \phi_2), \dots, (l_i, \phi_i), \dots, (l_n, \phi_n)\} \quad (15)$$

式中: 链路质量等级时间序列 $I_i = \{l_i, l_{i+1}, \dots, l_{i+w-1}\}$; $\phi_i = l_{i+w}$, ϕ_i 作为标签值。

滑动时间窗口的大小影响模型的预测结果。如果时间序列过短, 就会导致历史信息学习的缺失; 反之, 如果时间序列太长, 则会增加模型的复杂度, 甚至使预测效果变差。本文通过实验确定时间窗口大小。

3.2 基于 GRU 的链路质量预测

本文利用 GRU 能够很好地处理和预测序列数据

的优点, 构建链路质量预测模型 GRU_LQP (link quality prediction based on gate recurrent unit), 如图 1 所示。

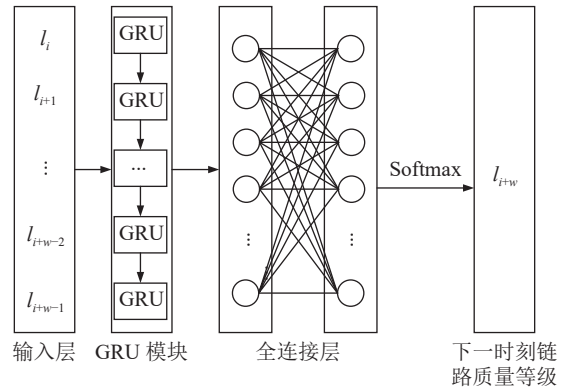


图 1 GRU_LQP 的结构

Fig. 1 Structure of GRU_LQP

由图 1 可知: GRU_LQP 包括输入层、GRU 模块层、全连接层和输出层。输入链路质量等级时间序列 $\{l_i, l_{i+1}, \dots, l_{i+w-1}\}$, 通过 GRU 模块中的 GRU 单元学习输入序列之间的时序关系, 通过全连接层集成链路质量等级时序特征。为了增强链路质量预测模型的泛化性, 在全连接层中加入 Dropout 操作, 设置其缩放权重为 0.5^[25]。选用 Softmax 函数映射得到下一时刻各等级的概率, 实现链路质量预测。

GRU 模块使用多个 GRU 单元的堆叠实现对输入链路质量等级时间序列的学习, GRU 单元的结构如图 2 所示。图 2 中, h_{t-1} 为前个 GRU 的隐藏层输出, l_t 为本轮 GRU 输入的链路质量等级, σ 为 Sigmoid 函数, z_t 和 r_t 分别为 GRU 的更新门和重置门, $\tilde{h}_t$ 为经过更新的中间状态, h_t 为本轮 GRU 的隐藏层输出。

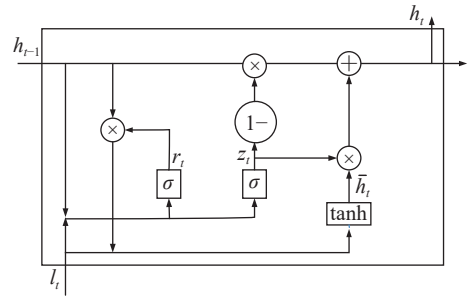


图 2 GRU 的结构

Fig. 2 Structure of GRU

GRU 主要通过两个门控对链路质量等级信息进行更新, 更新门 z_t 用于控制前一个状态中链路质量等级信息被代入到当前状态的程度, 更新门的值越接近 1, 说明代入了更多前一个状态的链路质量等级信息。重置门 r_t 控制前一状态有多少链路质量等级信息被写入到当前的候选集上, 重置门越小, 前一状态的

链路质量等级信息被写入得越少。通过Sigmoid函数将数据变换到0~1之间,并结合乘法操作可以实现对输入链路质量等级信息及历史链路质量等级信息的选择性重置。

GRU的计算过程如式(16)~(19)所示:

$$r_t = \sigma(w_r \cdot [l_t, h_{t-1}]) \quad (16)$$

$$z_t = \sigma(w_z \cdot [l_t, h_{t-1}]) \quad (17)$$

$$\bar{h}_t = \tanh(w_h \cdot [r_t \times h_{t-1}, l_t]) \quad (18)$$

$$h_t = (1 - z_t) \times h_{t-1} + z_t \times \bar{h}_t \quad (19)$$

式(16)~(19)中, w_r 为重置门的权重, w_z 为更新门的权重, w_h 为中间状态的权重。

本文使用随机梯度下降法优化权重参数 w_r 、 w_z 、 w_h , 学习率取经验值0.01。

3.3 预测模型的优化

为了进一步提高链路质量预测的准确率, 本文采用SVR作为链路质量等级预测器, 优化GRU_LQP, 构建了GRU_SVR(link quality prediction based on gate recurrent unit and support vector regression)链路质量预测模型, 如图3所示。

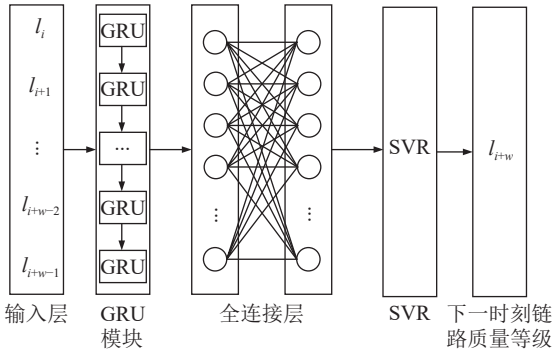


图3 GRU_SVR的结构

Fig. 3 Structure of GRU_SVR

GRU_SVR的参数设置如表1所示。

表1 GRU_SVR参数

Tab. 1 Parameters of GRU_SVR

参数属性	取值
全连接层层数	2
drop out	0.5
epoch	100
学习率	0.01
损失函数	均方误差损失函数

3.4 模型评价

本文采用均方误差(mean square error, MSE)评价预测结果, MSE的计算式为:

$$R_{\text{MSE}} = \frac{1}{n} \sum_{i=1}^n (\phi'_i - \phi_i)^2 \quad (20)$$

式中, ϕ_i 为样本中输入链路质量等级时间序列的标签, ϕ'_i 为预测值。MSE值越小则表明预测模型预测误差越小, 性能越好。

4 实验设计与结果分析

在3个真实场景中进行实验, 通过与同类模型对比验证本文模型的有效性。

4.1 实验设计

分别在实验室、走廊及停车场布置传感器节点, 如图4所示。实验室场景中, 无线设备及人员的走动等对链路质量造成影响; 走廊场景中, 节点产生的串扰及人员的走动等对链路质量造成影响; 停车场场景中, 车辆出入及车载设备等对链路质量造成影响。



(a) 实验室场景



(b) 走廊场景



(c) 停车场场景

图4 实验场景

Fig. 4 Experimental scenarios

实验在南昌航空大学物联网实验室开发的WSN链路质量测试平台(wireless sensor networks link quality testbed, WSN-LQT)^[25]上进行, 连续3天收集RSSI、LQI、SNR、PRR等链路质量数据, 将预处理后的链路质量样本集以7:3的比例划分为评估模型的训练集和测试集, 根据评估结果构建链路质量等级时序样本集, 并以7:3的比例划分为预测模型的训练集和测试集。WSN-LQT中参数设置如表2所示。

表 2 WSN-LQT参数

Tab. 2 Parameters of WSN-LQT

参数	参数值
信道	26
周期内探测包数	30
数据包发送间隔/s	0.2
测试周期/s	12
监听方式	主动监听
传输速度/kbps	250

4.2 主要实验参数确定

1) 聚类数目

采用Gap Statistic算法确定聚类数目 K , 采用K-means++聚类算法对PRR聚类, 得到链路质量等级。设 $\Gamma = Gap(K) - Gap(K + 1) + s_{K+1}$, 3个实验场景中不同 K 值对应的 Γ 值如图5所示。

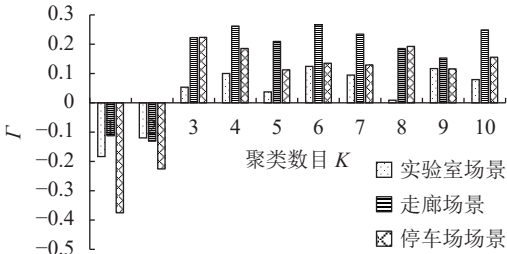


图 5 不同 K 值对应的 Γ 值

Fig. 5 Values of Γ corresponding to different K values

由图5可知, 3个实验场景中, 使式(5)成立的 K 值均为3。因此, 将实验室、走廊及停车场3个实验场景的PRR均划分为3个链路质量等级, 其中, 第1等级的链路为链路质量差的链路, 第2等级的链路为链路质量中等的链路, 第3等级的链路为链路质量好的链路。

2) 评估模型的参数

CatBoost_LQE评估模型的参数主要有: 学习率 lr 、L2正则参数 $l2_leaf_reg$ 、树的深度 $depth$ 、决策树最大的数量 $iterations$ 。为了提高CatBoost_LQE的评估准确率, 本文采用网格搜索法对上述参数寻优, 结果如表3所示。

表 3 CatBoost_LQE的参数值

Tab. 3 Parameter values of CatBoost_LQE

参数	depth	lr	l2_leaf_reg	iterations
实验室场景	5	0.03	9	300
走廊场景	4	0.03	3	500
停车场场景	4	0.10	4	300

3) 滑动时间窗口的大小

本文通过对比不同时间窗口下GRU_SVR的MSE, 选择合适的时间窗口大小。3个场景中, 不同的时间

窗口下GRU_SVR的MSE结果如图6所示。

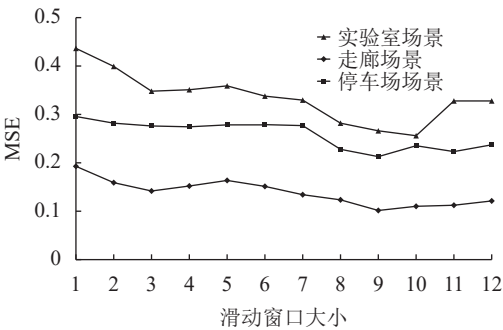


图 6 GRU_SVR的MSE

Fig. 6 MSE of GRU_SVR

由图6可知, 滑动时间窗口大小影响预测模型的准确率。在实验室场景中, 滑动时间窗口大小为10时, MSE值最小。在停车场和走廊场景中, 滑动时间窗口大小为9时, MSE值最小。因此, 实验室、走廊及停车场场景中, 滑动时间窗口的大小分别设置为10、9、9。

4.3 模型的验证

为验证GRU_SVR预测模型的有效性, 与GRU_LQP、4C^[9]、WNN_LQP^[10]、RFVL_LQP^[11]及RNN_LQP^[12]预测模型进行对比, 结果如图7所示。

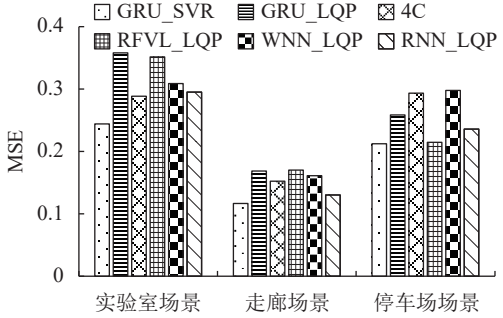


图 7 不同预测模型的MSE值

Fig. 7 MSE values of different prediction models

由图7可知: 实验室场景中, 干扰较多, 链路质量整体较差, GRU_SVR的均方误差最小。走廊实验场景中, 链路的干扰较少, 链路质量整体较好, GRU_SVR的均方误差也是最小的。可见, 与GRU_LQP、4C、WNN_LQP、RFVL_LQP及RNN_LQP预测模型相比, 本文提出的GRU_SVR预测模型无论是在链路质量较差还是较好的环境中, 均有更好的预测准确率。

5 结 论

本文提出一种基于GRU的链路质量预测方法。首先, 采用基于Gap Statistic算法优化的K-means++算法划分链路质量等级, 获得样本标签; 然后, 选择RSSI、LQI、SNR作为输入, 基于CatBoost构建链路质量评估模型, 评估当前时刻的链路质量, 在不同场景中使用网格搜索法优化评估模型的参数; 最后, 采用

GRU学习链路质量等级序列的时序信息,采用SVR预测下一时刻的链路质量等级。在真实世界中,部署节点并收集数据进行实验,结果表明相比于其他模型,本文所提模型具有更低的预测误差。下一步将致力于解决链路质量数据不平衡给链路质量预测带来的影响,进一步提高链路质量预测的准确率。

参考文献:

- [1] Singh S,Saini H S.Intelligent ad-hoc-on demand multipath distance vector for wormhole attack in clustered WSN[J]. [Wireless Personal Communications](#),2022,122(2):1305–1327.
- [2] Fascista A.Toward integrated large-scale environmental monitoring using WSN/UAV/crowdsensing:A review of applications,signal processing,and future perspectives[J]. [Sensors](#),2022,22(5):1824.
- [3] Feng Sheng,Shi Haiyan,Huang Longjun,et al.Unknown hostile environment-oriented autonomous WSN deployment using a mobile robot[J].[Journal of Network and Computer Applications](#),2021,182:103053.
- [4] Pan M S,Yeh L W,Chen Y A,et al.A WSN-based intelligent light control system considering user activities and profiles[J].[IEEE Sensors Journal](#),2008,8(10):1710–1721.
- [5] Wei Liansuo,Han Jian,Jin Tao,et al.Topology optimization algorithm based on optimal rigid sub-graph for potential-game wireless sensor networks[J].[Advanced Engineering Sciences](#),2021,53(2):125–132.[魏连锁,韩建,金涛,等.基于最优刚性子图的势博弈无线传感器网络拓扑优化算法[J].[工程科学与技术](#),2021,53(2):125–132.]
- [6] Miao Chunyu,Chen Lina,Wu Jianjun,et al.Node location verification framework for WSN[J].[Journal of Computer Research and Development](#),2019,56(6):1231–1243.[苗春雨,陈丽娜,吴建军,等.无线传感器网络节点位置验证框架[J].[计算机研究与发展](#),2019,56(6):1231–1243.]
- [7] 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):239–244.[翁丽娜,刘铁铭,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):239–244.]
- [8] Sharma A,Bansal A,Rishiwal V.Selection of high quality path through MXAODV in mobile ad-hoc network[J].[International Journal of Systems,Control and Communications](#),2019,10(1):1.
- [9] Liu Tao,Cerpa A E.Foresee (4C):Wireless link prediction using link features[C]//[Proceedings of the 10th ACM/IEEE International Conference on Information Processing in Sensor Networks](#).Piscataway:IEEE,2011:294–305.
- [10] Liu Wei,Xu Jinwei,Yu Xia,et al.Wavelet neural network based link quality prediction for fluctuating low power wireless links[C]//[Proceedings of the 2021 IEEE 6th International Conference on Computer and Communication Systems](#).Chengdu:IEEE,2021:899–904.
- [11] Xue Xue,Sun Wei,Wang Jianping,et al.RVFL-LQP:RVFL-based link quality prediction of wireless sensor networks in smart grid[J].[IEEE Access](#),2020,8:7829–7841.
- [12] Xu Ming,Xu Jinwei,Xia Yu,et al.Recurrent neural network based link quality prediction for wireless sensor networks[C]//[Proceedings of the 2021 IEEE 6th International Conference on Computer and Communication Systems](#).Chengdu:IEEE,2021:905–909.
- [13] Zhao Rui,Wang Dongzhe,Yan Ruqiang,et al.Machine health monitoring using local feature-based gated recurrent unit networks[J].[IEEE Transactions on Industrial Electronics](#),2018,65(2):1539–1548.
- [14] Gou Chengcheng,Qin Yujun,Tian Tian,et al.Social messages outbreak prediction model based on recurrent neural network[J].[Journal of Software](#),2017,28(11):3030–3042.[简程成,秦宇君,田甜,等.一种基于RNN的社交消息爆发预测模型[J].[软件学报](#),2017,28(11):3030–3042.]
- [15] Liu Jiexi,Chen Songcan.Non-stationary multivariate time series prediction with MIX gated unit[J].[Journal of Computer Research and Development](#),2019,56(8):1642–1651.[刘颀羲,陈松灿.基于混合门单元的非平稳时间序列预测[J].[计算机研究与发展](#),2019,56(8):1642–1651.]
- [16] Abdel-Nasser M,Mahmoud K,Omer O A,et al.Link quality prediction in wireless community networks using deep recurrent neural networks[J].[Alexandria Engineering Journal](#),2020,59(5):3531–3543.
- [17] Ribeiro R G,Rios R.Temporal gap statistic:A new internal index to validate time series clustering[J].[Chaos,Solitons & Fractals](#),2021,142:110326.
- [18] Mydhili S K,Periyanyagi S,Baskar S,et al.Machine learning based multi scale parallel K-means++ clustering for cloud assisted Internet of Things[J].[Peer-to-Peer Networking and Applications](#),2020,13(6):2023–2035.
- [19] Yang M S,Sinaga K P.A feature-reduction multi-view k-means clustering algorithm[J].[IEEE Access](#),2019,7:114472–

114486.

- [20] Huang Xiaohui,Wang Cheng,Xiong Liyan,et al.A weighting k-means clustering approach by integrating intra-cluster and inter-cluster distances[J].[Chinese Journal of Computers](#), 2019,42(12):2836–2848.[黄晓辉,王成,熊李艳,等.一种集成簇内和簇间距离的加权k-means聚类方法[J].[计算机学报](#),2019,42(12):2836–2848.]
- [21] Sinaga K P,Yang M S.Unsupervised K-means clustering algorithm[J].[IEEE Access](#),2020,8:80716–80727.
- [22] Jiang Qigang,Yang Xiuyan,Yang Changbao,et al.Object-oriented land use classification based on CatBoost algorithm[J].[Journal of Jilin University\(Information Science Edition\)](#), 2020,38(2):185–191.[姜琦刚,杨秀艳,杨长保,等.基于CatBoost算法的面向对象土地利用分类[J].[吉林大学学报\(信息科学版\)](#),2020,38(2):185–191.]
- [23] Hancock J T,Khoshgoftaar T M.CatBoost for big data:An interdisciplinary review[J].[Journal of Big Data](#),2020,7:94.
- [24] Samat A,Li Erzhu,Du Peijun,et al.CatBoost for RS image classification with pseudo label support from neighbor patches-based clustering[J].[IEEE Geoscience and Remote Sensing Letters](#),2022,19:1–5.
- [25] Shu Jian,Gao Su,Chen Yubin.Link quality estimator based on adaptive general regression neural network[J].[Journal of Computer Research and Development](#),2020,57(12):2662–2672.[舒坚,高素,陈宇斌.基于自适应广义回归神经网络的链路质量评估[J].[计算机研究与发展](#),2020,57(12):2662–2672.]

(编辑 赵 婧)

引用格式: Liu Linlan,Xiao Tingzhong,Shu Jian,et al.Link quality prediction based on gate recurrent unit[J].[Advanced Engineering Sciences](#),2022,54(6):51–58.[刘琳岚,肖庭忠,舒坚,等.基于门控循环单元的链路质量预测[J].[工程科学与技术](#),2022,54(6):51–58.]