

基于高斯过程回归的链路质量预测模型

舒坚¹, 刘满兰², 尚亚青², 陈宇斌¹, 刘琳岚²

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

摘 要: 基于链路质量的路由选择机制可有效感知当前链路的变化, 且对无线传感器网络的可靠通信起着重要作用, 基于此, 提出基于高斯过程回归的链路质量预测模型。通过灰关联方法计算链路质量参数与分组接收率的关联度, 选取链路质量指示均值和信噪比均值作为模型的输入参数, 以降低计算复杂度。采用链路质量指示均值、信噪比均值和分组接收率构建基于组合协方差函数的高斯过程回归模型预测链路质量。稳定场景与不稳定场景下的实验结果表明, 与动态贝叶斯网络预测模型相比, 所提模型具有更好的预测精确度。

关键词: 无线传感器网络; 高斯过程回归; 链路质量预测; 组合协方差函数; 灰关联算法

中图分类号: TP393

文献标识码: A

doi: 10.11959/j.issn.1000-436x.2018113

Link quality prediction model based on Gaussian process regression

SHU Jian¹, LIU Manlan², SHANG Yaqing², CHEN Yubin¹, LIU Linlan²

1. School of Software, Nanchang Hangkong University, Nanchang 330063, China

2. School of Information Engineering, Nanchang Hangkong University, Nanchang 330063, China

Abstract: Link quality is an important factor of reliable communication and the foundation of upper protocol design for wireless sensor network. Based on this, a link quality prediction model based on Gaussian process regression was proposed. It employed grey correlation algorithm to analyze correlation between link quality parameters and packet receive rate. The mean of the link quality indication and the mean of the signal-to-noise were selected as input parameters so as to reduce the computational complexity. The above parameters and packet receive rate were taken to build Gaussian process regression model with combination of covariance function, so that link quality could be predicted. In the stable and unstable scenarios, the experimental results show that the proposed model has better prediction accuracy than the one of dynamic Bayesian network prediction model.

Key words: wireless sensor network, Gaussian process regression, link quality prediction, combination of covariance function, grey correlation algorithm

1 引言

近年来, 随着微机电、无线通信和嵌入式等技术的不断发展和相互融合, 低成本、低功耗、多功能的微型传感器得到了广泛应用。

无线传感器网络(WSN, wireless sensor network)是一种由大量廉价微型传感器构成、具有动态拓扑结构、以数据传输为中心的多跳自组织网络^[1]。节点部署环境的复杂多变导致数据传输过程易受周围环境的干扰和影响, 链路特性表现为不规则性、非

收稿日期: 2017-09-18; 修回日期: 2018-05-25

通信作者: 刘琳岚, liulinlan@nchu.edu.cn

基金项目: 国家自然科学基金资助项目(No.61762065, No.61363015, No.61501218, No.61501217); 江西省自然科学基金资助项目(No.20171BAB202009, No.20171ACB20018)

Foundation Items: The National Natural Science Foundation of China (No.61762065, No.61363015, No.61501218, No.61501217), The Natural Science Foundation of Jiangxi Province (No.20171BAB202009, No.20171ACB20018)

对称性和时空波动性^[2-4]。链路质量对无线传感器网络数据吞吐量、传输时延和能耗等性能具有至关重要的作用^[2]。有效的链路质量预测机制可以通过在邻居列表中选择高质量的链路进行通信，提高数据传输的可靠性^[5]。

本文在分析现有链路质量预测方法的基础上，采用高斯过程回归（GPR, Gaussian process regression）方法建立链路质量预测模型。本文的主要贡献如下。

1) 通过灰关联分析法对分组接收率与接收信号强度指示均值、链路质量指示均值、信噪比均值之间的灰关联度进行分析计算，并选取其中与分组接收率关联度大的链路质量指示均值和信噪比均值作为本文的链路质量参数。

2) 结合 SE 协方差函数和 Matern 协方差函数，构建基于组合协方差函数的高斯过程回归模型 C-GPR，采用标准误差评价预测的精确度。

2 相关研究

2.1 链路质量预测

通过基于链路质量的路由选择机制，能有效感知当前链路的变化状况，选择一个最好的转发者为当前节点的下一跳，可有效减少数据重发次数。现有的链路质量预测方法包括基于硬件的链路质量预测方法、基于软件的链路质量预测方法、综合性的链路质量预测方法以及基于机器学习理论的链路质量预测方法。

1) 基于硬件的链路质量预测方法主要涉及接收信号强度指示（RSSI, received signal strength indication）、链路质量指示（LQI, link quality indication）、信噪比（SNR, signal-to-noise）等物理层参数。文献[1]提出 RSSI 和 LQI 均值测量方法，通过测量和计算 RSSI 和 LQI 均值获得链路质量信息，若存在多个干扰，该方法无法可靠地反映链路质量状况。文献[6]指出当 RSSI 值超过 -87 dBm 时，分组接收率（PRR, packet receive rate）一般高于 85%；而当 RSSI 低于 -87 dBm 时，链路质量进入灰色区域。文献[7]指出 LQI 相对于 RSSI 能够更直接地展现信号质量，但 LQI 仅为测量成功接收数据分组的平均相关性，没有考虑数据分组的丢失情况。文献[8]采用高阶卡尔曼滤波器对 SNR 进行处理，提出 Effective-SNR 链路质量指标，实验结果表明该指标

与 PRR 相关性较高。上述研究表明，基于硬件的链路质量预测方法通过直接从硬件中获得物理层参数进行预测，计算量小、资源消耗少且对链路变化反映较灵敏，但单个物理层参数不能全面反映链路质量的变化。

2) 基于软件的链路质量预测方法主要包括 PRR、ETX 等。文献[9]采用滑动窗口加权指数均值（WMEWMA, window mean exponentially weighted moving average）法计算 PRR 的值，通过调整时间窗口 w 和权值系数 α ，在稳定性和灵敏度之间获得良好的权衡。文献[10]提出自适应的链路质量估计方法 EasiLQE，采用基于误差的滤波器估计 PRR 值，实验表明，在复杂的网络环境下该方法比 WMEWMA 准确性更好、能量消耗更低，适用于突发较频繁的链路。文献[11]基于链路的分组丢失率和不对称性提出期望分组传输次数（ETX, expected transmission count）指标，其通过计算端到端正反向链路 PRR 得到。ETX 作为链路质量度量标准，在汇聚树协议（CTP, convergence tree protocol）和动态原路由（DSR, dynamic source routing）协议中广泛使用。相对物理层参数，上述方法只基于某一段时间内的 PRR 值，因此不能快速地反映无线链路的频繁变化。

3) 综合性的链路质量预测方法融合了物理层参数和链路层参数。文献[12]结合 PRR、LQI 及 SNR 参数相关性，采用 LQI 和 SNR 的均值构建三角度量（TM, triangle metric）指标 $d_{\Delta} = \sqrt{\overline{SNR}_w^2 + \overline{LQI}_w^2}$ ，通过收集 10 个数据分组预测动态变化的链路质量，但没有考虑分组丢失的情况。文献[13]将 PRR、链路稳定性因子、非对称性以及信噪比均值等指标通过模糊规则计算它们的综合值，提出基于模糊规则的链路质量预测（F-LQE, fuzzy-based link quality estimation）方法，实验结果表明，这种综合性的预测方法比单一指标的预测方法具有更高的准确率，并在稳定性和灵敏度上优于 ETX 和 WMEWMA。相比前 2 种方法，这类方法能够较全面地反映链路质量，但能量消耗较大。

4) 基于机器学习理论的链路质量预测方法通过机器学习算法对链路质量参数建模进行预测。文献[14]提出基于逻辑回归的链路质量预测机制，以 RSSI、LQI 和 SNR 为输入，下一时刻 PRR 为输出，

室内、室外的实验结果表明,该方法能够节省 20%~30% 的传输成本。文献[15]提出模糊支持向量回归机链路质量预测模型,首先采用无监督模糊核聚类算法自动划分样本集获取样本隶属度,然后采用混沌粒子群优化算法选择子模型参数,实验结果表明,该模型比 BP 神经网络预测模型具有更好的预测精度。文献[16]采用非参数时间序列对链路质量参数建模,提出基于模型框架的链路质量预测(N-STLP, nonparametric short-term link prediction)方法,结果表明, N-STLP 具有较高的精确度和收敛速度。文献[17]提出基于动态贝叶斯网络(DBN, dynamic Bayesian network)的链路质量预测方法,通过非对称贴适度构建综合性的链路质量等级指标,采用最大期望算法进行 DBN 模型的参数学习,具有较好的预测准确度。相比于传统数学统计方法,机器学习方法在处理复杂关系的非线性数据建模问题上具有优势^[18]。

2.2 高斯过程回归

WSN 部署环境复杂,链路质量的变化随自身校准以及内部干扰和外部因素影响而变化,由多种因素共同决定,因此链路质量预测为复杂非线性问题。GPR 是一种基于函数分布的机器学习算法,提供了基于贝叶斯框架的无参数核方法,在处理小样本、高维、线性不可分等复杂问题上具有优势^[19]。文献[20]采用整体经验模式分解(EEMD, ensemble empirical mode decomposition),提出基于高斯过程回归的飞机辅助动力装置退化信号的概率预测方法,与 GPR 相比, EEMD-GPR 方法获得了较高的精度。文献[21]采用 GPR 对具有不可控性和随机性的短期风速进行预测。文献[22]针对网络流量的不确定性、突发性和混沌性等特性,提出基于 GPR 的网络流量预测模型,并取得了良好的预测效果。文献[23]提出基于高斯过程模型的多步预测方法,实现了对未来时刻环境监测数据的预测,该模型可用于环境参数未来趋势分析、异常环境事件预警等场景。文献[24]针对混沌时间序列单步和多步预测,利用矩阵运算和优化算法自适应地确定协方差函数和均值函数中的超参数,提出基于复合协方差函数的高斯过程模型,应用该模型能较精确地预测混沌时间序列,且具有稳健性。文献[25]采用高斯过程回归的概率预测模型对区域电网进行饱和负荷预测,通过改进混沌粒子群算法优化模型的超参数,在多个情景下均具有较高的预测精度,并且可

增强其预测的弹性。文献[26]采用高斯混合模型预测复杂运动对象的运动轨迹,大大缩减了其轨迹预测的时间。文献[27]提出基于人工蜂群优化的高斯过程分类方法,用于脑电信号的模式识别。

本文采用 GPR 预测链路质量,考虑到如果选取过多的物理层参数作为预测模型的输入,参数间具有一定的相关性,涵盖的链路质量信息存在冗余,势必导致模型结构复杂、计算量增加,因此,采用灰关联分析法量化 PRR 与各物理层参数之间的灰关联度,筛选出影响链路质量的参数集作为预测模型的输入。为提高模型预测精确度,结合平方指数(SE, squared exponential)协方差函数和 Matern 协方差函数作为 GPR 模型的协方差函数,构建基于组合协方差函数的高斯过程回归模型 C-GPR。

3 链路质量参数的选取

3.1 链路质量参数的相关性

文献[28]指出基于 CC2420 的传感器节点的接收功率高于阈值(-87 dBm)时, RSSI 与 PRR 呈强相关性,而 LQI 更能反映中等质量链路的变化。文献[29]也证实,当接收功率低于节点阈值时, RSSI 与分组丢失率之间无特定关系。设置含有背景噪声、信道干扰及多径衰减环境因素的 3 个实验场景,结果表明,室内环境中均值 LQI 与 PRR 之间的 Pearson 系数大于 RSSI 与 PRR 的 Pearson 系数。

文献[7]指出连通区中 RSSI 和 LQI 均能指示链路质量状况,而在灰色区域中,均值 LQI 能更好地反映链路质量状况。文献[30]指出,物理层参数直接从硬件中取出,对链路质量变化状况的反映灵敏,当 SNR 值较低时,误比特率(BER, bite error rate)增加,动态网络中 SNR 值更能灵活反映邻近节点链路质量的变化情况。

上述研究表明,物理层参数与 PRR 之间存在相关性,均值化的参数相比单个参数更精确。当链路质量较高时,各物理层参数之间存在强相关性,对于中等链路,各参数时间序列变化趋势不一致。

3.2 采用灰关联分析法选取链路质量参数

通过灰关联分析法找出影响系统行为的主次元素^[31],系统各因子行为序列变化趋势的接近程度反映了相应序列之间的关联度^[32]。为了消除单位不同带来的影响,在计算灰关联度之前,需要对原始数据进行无量纲化。

本文分别获取稳定和不稳定这 2 种场景下的 RSSI 均值 \overline{RSSI} 、LQI 均值 \overline{LQI} 、SNR 均值 \overline{SNR} 和 PRR，并分别通过式(1)~式(3)计算。采用灰关联分析法对 PRR 与这 3 个参数之间的灰关联度进行排序，筛选出关联度大的链路质量参数集作为模型的输入。

$$\overline{RSSI} = \frac{\sum_{i=1}^m RSSI_i}{m} \quad (1)$$

$$\overline{LQI} = \frac{\sum_{i=1}^m LQI_i}{m} \quad (2)$$

$$\overline{SNR} = \frac{\sum_{i=1}^m SNR_i}{m} \quad (3)$$

其中， m 为通信周期内接收的数据分组的数量。假设收集了 s 个周期的数据，则可得 RSSI 向量 $\overline{RSSI} = (\overline{RSSI}_1, \overline{RSSI}_2, \dots, \overline{RSSI}_s)$ 、LQI 向量 $\overline{LQI} = (\overline{LQI}_1, \overline{LQI}_2, \dots, \overline{LQI}_s)$ 、SNR 向量 $\overline{SNR} = (\overline{SNR}_1, \overline{SNR}_2, \dots, \overline{SNR}_s)$ 和 PRR 向量 $\overline{PRR} = (PRR_1, PRR_2, \dots, PRR_s)$ 。

设 $\mathbf{X} = \{\mathbf{X}_0, \mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3\}$ 为灰关联因子集，其中， \mathbf{X}_0 为 PRR，代表系统变量，有 k 个值， \mathbf{X}_1 为 \overline{RSSI} ， \mathbf{X}_2 为 \overline{LQI} ， \mathbf{X}_3 为 \overline{SNR} ，有 K 个观测数据为 $x_i(k) (i=1,2,3; k=1,2,\dots,K)$ 。

灰关联度的计算步骤描述如下。

Step1 各序列数据无量纲化。采用式(4)对 \mathbf{X}_i 进行无量纲化，得到 \mathbf{X}'_i ，即

$$\begin{aligned} X'_i &= \frac{X_i}{x_i(1)} \\ &= (x'_i(1), x'_i(2), \dots, x'_i(K)), i=0,1,2,3 \end{aligned} \quad (4)$$

Step2 采用式(5)和式(6)求差序列

$$\Delta_i(k) = |x'_0(k) - x'_i(k)| \quad (5)$$

$$\Delta_i = (\Delta_i(1), \Delta_i(2), \dots, \Delta_i(K)), i=1,2,3 \quad (6)$$

Step3 采用式(7)和式(8)分别求两级最大差 M 和最小差 m

$$M = \max_i \max_k \Delta_i(k) \quad (7)$$

$$m = \min_i \min_k \Delta_i(k) \quad (8)$$

Step4 采用式(9)求关联系数 $\gamma_{oi}(k)$ ，即

$$\gamma_{oi}(k) = \frac{m + \xi M}{\Delta_i(k) + \xi M}, i=1,2,3; k=1,2,\dots,K \quad (9)$$

其中， ξ 为分辨系数， $\xi \in (0,1)$ ，本文取经验值为 0.5。

Step5 采用式(10)计算关联度 γ_{oi} ，即

$$\gamma_{oi} = \frac{1}{K} \sum_{k=1}^K \gamma_{oi}(k), i=1,2,3 \quad (10)$$

其中， γ_{01} 、 γ_{02} 、 γ_{03} 分别表示 PRR 与 \overline{RSSI} 、PRR 与 \overline{LQI} 、PRR 与 \overline{SNR} 的灰关联度。通过数据处理系统软件计算 γ_{01} 、 γ_{02} 和 γ_{03} 的值，选取灰关联度较大的 \overline{LQI} 和 \overline{SNR} 作为预测模型的输入参数。

4 链路质量预测

在高斯过程回归模型中，协方差函数 \mathbf{K} 为半正定矩阵，满足 Mercer 条件，故协方差函数等价于核函数。协方差函数影响模型预测的精确度，其选取取决于样本数据的分布和样本输入与输出之间的关系^[33]。本文结合 SE 协方差函数和 Matern 协方差函数，构建 C-GPR 链路质量预测模型，采用标准误差 (RMSE, root mean squared error)^[34] 评价预测模型的精确度。

4.1 协方差函数的选取

SE 协方差函数通过计算 2 个数据点之间的空间距离得到协方差的值，插值能力比较强，Matern 协方差函数能对短期不规则性的时间序列数据进行良好的拟合。

1) SE 协方差函数

$$C_{SE}(\boldsymbol{\varphi}_i, \boldsymbol{\psi}_j) = \sigma_f^2 \exp\left(-\frac{(\boldsymbol{\varphi}_i - \boldsymbol{\psi}_j)^2}{2q^2}\right) + \sigma_n^2 \delta_{ij} \quad (11)$$

其中， $\boldsymbol{\varphi}_i$ 为 \overline{LQI} ， $\boldsymbol{\psi}_j$ 为 \overline{SNR} 。 δ_{ij} 为 Kronecker 算子，若 $i=j$ ，那么 $\delta_{ij}=1$ ，否则 $\delta_{ij}=0$ 。 σ_f^2 为 SE 协方差函数的信号方差，通过它能改变数据点间的局部关联性， σ_n^2 为噪声的方差， q 为超参数，用来计算数据点间的关联程度， q 值越小，相关程度越大。

2) Matern 协方差函数

$$C_{\text{Matern}}(r) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu}r}{q}\right)^\nu K_\nu\left(\frac{\sqrt{2\nu}r}{q}\right) \quad (12)$$

其中, $r = |\varphi_i - \varphi_j|$, $v = \frac{p+1}{2}$, p 为正数, $\theta = \{\sigma_f^2, \sigma_n^2, q\}$ 。对于多维输入变量, 各向同性平稳协方差函数通过设置 $r^2(\varphi_i - \varphi_j) = (\varphi_i - \varphi_j) \mathbf{M}(\varphi_i - \varphi_j)$ 改变输入变量每个维度的特征尺度, 实现不同维度的各向异性, 其中, \mathbf{M} 为自动关联测定 (ARD, automatic relevance determination) 的半正定对角矩阵。

4.2 超参数的优化

本文提出的 GPR 链路质量预测模型中组合协方差函数的超参数 θ 通过极大似然法进行优化, 其训练样本的负对数边缘似然函数为

$$L = -\lg P(\mathbf{y} | \mathbf{Z}, \theta) = \frac{1}{2} \mathbf{y}^T \mathbf{K}^{-1} \mathbf{y} + \frac{1}{2} \lg |\mathbf{K}| + \frac{n}{2} \lg(2\pi) \quad (13)$$

$$\frac{\partial L}{\partial \theta_j} = \frac{1}{2} \mathbf{y}^T \mathbf{K}^{-1} \frac{\partial \mathbf{K}}{\partial \theta_j} \mathbf{K}^{-1} \mathbf{y} + \frac{1}{2} \text{tr} \left(\mathbf{K}^{-1} \frac{\partial \mathbf{K}}{\partial \theta_j} \right) \quad (14)$$

其中, tr 表示矩阵的迹, \mathbf{Z} 表示模型的输入向量, \mathbf{y} 表示模型的输出向量。在极大似然框架下, 计算 $\lg[p(\mathbf{y} | \mathbf{Z}, \theta)]$ 对 θ 的偏导数, 并令偏导数为 0, 采用共轭梯度算法优化超参数 θ 的值, 得到对数边缘似然函数的最大值。

4.3 C-GPR 模型

高斯过程全部统计特征由其均值 $\mu(z)$ 及协方差 $C(z, z')$ 确定, 定义为

$$f(z) \sim GP\{\mu(z), C(z, z')\} \quad (15)$$

设数据集 $\mathbf{D} = \{(\mathbf{Z}_i, \mathbf{y}_i) | i = 1, 2, \dots, n\}$ 为高斯过程回归模型的训练集, 其中, $\mathbf{Z}_i = (\overline{LQI}_i, \overline{SNR}_i)$ 为预测模型的输入, $\mathbf{y} = (y_1, y_2, \dots, y_n)^T$ 为 \mathbf{PRR} 向量集, 是预测模型的输出。带有高斯白噪声的标准线性回归模型为

$$\mathbf{y} = f(\mathbf{Z}) + \varepsilon \quad (16)$$

$$f(\mathbf{Z}) = \phi(\mathbf{Z})^T \mathbf{w} + b \quad (17)$$

其中, \mathbf{Z} 为输入向量, f 为函数值, \mathbf{y} 为受加性噪声污染的观测值, ε 为独立的高斯白噪声, 符合均值为 0、方差为 σ_n^2 的高斯分布, 即 $\varepsilon \sim N(0, \sigma_n^2)$, 由于噪声 ε 为独立于 $f(\mathbf{Z})$ 的白噪声, 当 $f(\mathbf{Z})$ 服从高斯分布时, \mathbf{y} 同样服从高斯分布, 假设 \mathbf{w} 的先验概率为 $\mathbf{w} \sim N(0, \mathbf{I})$, \mathbf{I} 为单位矩阵, 则 $\mathbf{y} \sim N(0, \mathbf{C} + \sigma_n^2 \mathbf{I})$ 。对于测试输入 $\mathbf{Z}^* = (\overline{LQI}^*, \overline{SNR}^*)$, 根据 GP 中 \mathbf{y} 的

先验知识, 建立样本输出和测试样本输出 \mathbf{y}^* 所形成的联合高斯先验分布为

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{y}^* \end{bmatrix} \sim N \left(0, \begin{bmatrix} \mathbf{C}(\mathbf{Z}, \mathbf{Z}) + \sigma_n^2 \mathbf{I} & \mathbf{C}(\mathbf{Z}, \mathbf{Z}^*) \\ \mathbf{C}(\mathbf{Z}, \mathbf{Z}^*) & \mathbf{C}(\mathbf{Z}^*, \mathbf{Z}^*) \end{bmatrix} \right) \quad (18)$$

其中, $\mathbf{C}(\mathbf{Z}, \mathbf{Z})$ 为 $n \times n$ 对称正定协方差矩阵, c^i 度量了 \mathbf{Z}_i 和 \mathbf{Z}_j 的相关性, $\mathbf{C}(\mathbf{Z}, \mathbf{Z}^*)$ 为训练集 \mathbf{Z} 与测试点 \mathbf{Z}^* 的 $n \times 1$ 阶协方差矩阵, $\mathbf{C}(\mathbf{Z}^*, \mathbf{Z}^*)$ 为测试点 \mathbf{Z}^* 自身协方差。

本文的高斯过程回归模型如图 1 所示。在给定测试点 $(\overline{LQI}^*, \overline{SNR}^*)$ 和训练数据集 \mathbf{Z} 的条件下, 根据贝叶斯基本理论计算出 $p(\mathbf{y}^* | \mathbf{Z}^*, \mathbf{D})$, 则 \mathbf{y}^* 的后验分布为 $p(\mathbf{y}^* | \mathbf{Z}^*) \sim N(\hat{\mathbf{y}}(\mathbf{Z}^*), \hat{\sigma}(\mathbf{Z}^*))$, 其中, \mathbf{y}^* 的均值和方差分别为

$$\hat{\mathbf{y}}(\mathbf{Z}^*) = \mathbf{k}^T(\mathbf{Z}^*) (\mathbf{K} + \sigma_n^2 \mathbf{I})^{-1} \mathbf{y} \quad (19)$$

$$\hat{\sigma}(\mathbf{Z}^*) = \mathbf{k}(\mathbf{Z}^*, \mathbf{Z}^*) - \mathbf{k}^T(\mathbf{Z}^*) (\mathbf{K} + \sigma_n^2 \mathbf{I})^{-1} \mathbf{k}(\mathbf{Z}^*) \quad (20)$$

其中, \mathbf{k} 为核函数, \mathbf{K} 为半正定矩阵, 均值 $\hat{\mathbf{y}}(\mathbf{Z}^*)$ 为预测值输出。令 $h(\mathbf{Z}^*) = \mathbf{k}^T(\mathbf{Z}^*) (\mathbf{K} + \sigma_n^2 \mathbf{I})^{-1}$, 则 $\hat{\mathbf{y}}(\mathbf{Z}^*) = h(\mathbf{Z}^*) \mathbf{y}$, 均值 $\hat{\mathbf{y}}(\mathbf{Z}^*)$ 的预测可以看作目标值 \mathbf{y} 的线性组合, 通过核函数将其映射到线性特征空间, 从而转化成线性可分问题。

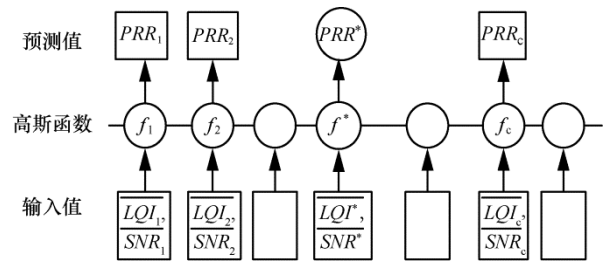


图 1 高斯过程回归模型

4.4 评价指标

本文从采集的数据样本中随机抽出 75% 的样本作为训练初始模型的训练集, 剩下的样本作为验证 GPR 预测模型精确度的测试集。采用标准误差评价预测精确度, 即

$$RMSE = \sqrt{\frac{1}{t} \sum_{i=1}^t (y_i - \hat{y}_i)^2} \quad (21)$$

其中, y_i 为真实值, \hat{y}_i 为预测值, t 为测试样本总数。

5 实验结果与分析

本文采用美国 CrossBow 公司研制的 TelosB 节点和南昌航空大学物联网技术研究所开发的链路质量测试平台 (LQT, link quality testbed) 分别对 PRR 与 \overline{RSSI} 、 \overline{LQI} 、 \overline{SNR} 的相关性, 链路质量参数 \overline{LQI} 、 \overline{SNR} 的有效性以及组合协方差函数的有效性进行验证, 并与 DBN 预测模型进行了对比实验。平台采用 26 信道通信, 功率设置为 0 dBm, 每 6 s 发送 30 个数据分组, 间隔为 200 ms, 空闲 4 s。

5.1 实验场景

本文在 2 个实验场景下以不同距离和方向均部署了 9 个节点组成的网络进行实验, 实验场景分别为树林和交通主干道, 8 个节点为感知节点, 一个节点为 Sink 节点。

树林实验场景如图 2 所示, Sink 节点被放置在中间, 在 Sink 节点东、西、南、北 4 个方向以不同的距离分别在每个方向部署 2 个节点。实验过程中, 由于节点各个方向的植物和石头等的分布情况不同, 导致各个方向的节点受到的反射、折射程度不同。但是树林中鲜少有人走动, 且障碍物处于静态, 影响信号传输的因素主要为多径效应, 不受其他信号及外界的干扰, 因此树林实验场景为稳定场景。

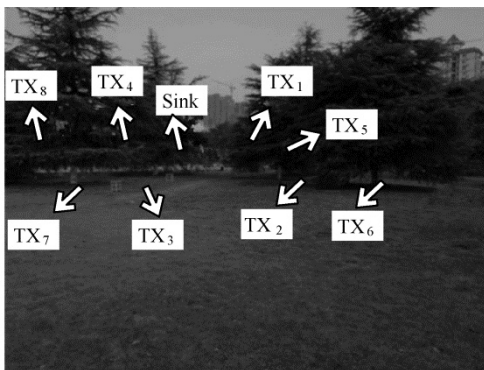


图 2 树林实验场景

公路实验场景如图 3 所示, 该路段为交通主干道, Sink 节点部署在绿化带中央, 道路两旁分别依次部署 4 个节点。实验时间为放学高峰期, 路上的车辆和行人较多, 他们的遮挡以及他们随身携带的电子设备的信号干扰都会对 WSN 的链路质量产生严重的影响, 因此公路实验场景为不稳定场景。

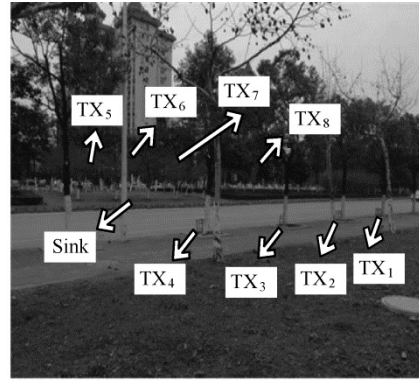


图 3 公路实验场景

5.2 PRR 与 \overline{RSSI} 、 \overline{LQI} 、 \overline{SNR} 的相关性

本文采用国产数据处理系统 (DPS, data processing system) 分别对随机选取的稳定场景和不稳定场景下的 200 个 PRR 、 \overline{RSSI} 、 \overline{LQI} 和 \overline{SNR} 序列进行灰关联分析, 得到稳定场景下和不稳定场景下的灰关联度值分别如表 1 和表 2 所示。

表 1 稳定场景下灰关联度值

参数	取值
γ_{01}	0.84
γ_{02}	0.89
γ_{03}	0.85

表 2 不稳定场景下灰关联度值

参数	取值
γ_{01}	0.68
γ_{02}	0.77
γ_{03}	0.71

由表 1 和表 2 可知, γ_{02} 的值最大, 这表明 PRR 与 \overline{LQI} 的灰关联度最高, 而 \overline{SNR} 与 PRR 的灰关联度略高于 \overline{RSSI} 与 PRR 的灰关联度, 且 \overline{RSSI} 与 \overline{SNR} 具有强相关性^[2], 因此选取 \overline{LQI} 和 \overline{SNR} 为链路质量参数。

5.3 链路质量参数 \overline{LQI} 、 \overline{SNR} 的有效性

采用不同样本得到 2 种场景下 C-GPR 模型的预测结果分别如图 4 和图 5 所示。其中, 3-C-GPR 为采用原始样本 ($\overline{RSSI}, \overline{LQI}, \overline{SNR}$) 的预测结果, C-GPR 为采用降维后样本 ($\overline{LQI}, \overline{SNR}$) 的预测结果, 降维前后的 RMSE 对比如表 3 所示。

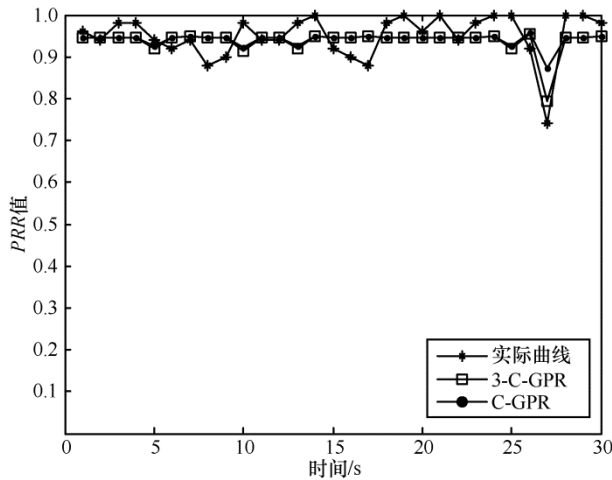


图 4 稳定场景下降维前后预测结果对比

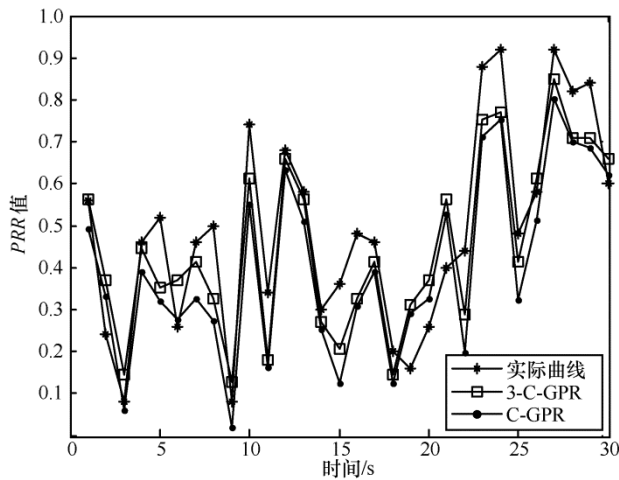


图 5 不稳定场景下降维前后预测结果对比

表 3 降维前后的 RMSE 对比

预测模型	稳定场景	不稳定场景
C-GPR	0.043 1	0.121 4
3-C-GPR	0.045 1	0.144 7

由图 4 和图 5 可知，2 种不同场景下样本降维前后的链路质量预测结果十分接近，均能准确地捕捉到链路质量的动态变化，且从表 3 可以看出，稳定和不确定场景下降维后的 C-GPR 预测精确度均高于 3-C-GPR。模型的计算复杂度为 $O(n^3m)$ ，其中， n 表示模型输入个数， m 表示梯度计算次数，而 3-C-GPR 模型输入维数多于 C-GPR，因此 3-C-GPR 所需存储空间较大，计算量也较大。这说明通过灰关联方法选取出的链路质量参数能够在保证预测精确度的同时，减少所需的存储空间，降低模型的计算量。

5.4 组合协方差函数的有效性

采用不同协方差函数得到 2 种场景下的预测 RMSE 如表 4 所示。SE-GPR 为采用单一 SE 协方差函数的预测模型，C-GPR 为采用组合协方差函数的预测模型。

表 4 不同协方差函数的 RMSE

预测模型	稳定场景	不稳定场景
C-GPR	0.043 1	0.121 4
SE-GPR	0.056 4	0.149 6

由表 4 可知，在稳定和不确定 2 种场景下，C-GPR 的 RMSE 分别为 0.043 1 和 0.121 4，而 SE-GPR 的 RMSE 分别为 0.056 4 和 0.149 6，均高于 C-GPR，这表明组合形式的协方差函数降低了模型的 RMSE，更有优势。在稳定场景下，由于链路波动较小，链路呈现较高的相关性，2 种模型的 RMSE 相差并不大，而在波动较大的不稳定场景下，2 种模型的 RMSE 差别显著，相比于单一协方差函数的 SE-GPR，C-GPR 在这种不规则的链路中具有更好的插值能力和拟合效果。

5.5 与 DBN 模型的对比

为进一步验证 C-GPR 模型的预测性能，采用文献[17]提出的 DBN 模型预测 2 种不同场景下的链路质量，得到 RMSE 如表 5 所示。

表 5 2 种预测模型的 RMSE

预测模型	稳定场景	不稳定场景
C-GPR	0.043 1	0.121 4
DBN	0.094 3	0.323 4

由表 5 可知，在稳定场景与不稳定场景下，C-GPR 的 RMSE 分别为 0.043 1 和 0.121 4，相比 DBN 模型，预测精确度有明显提高，表明本文提出的模型具有更好的预测精确度。

6 结束语

本文在分析现有链路质量预测方法的基础上，提出基于高斯过程回归的链路质量预测模型 C-GPR。

稳定场景与不稳定场景下的实验验证了链路质量参数 \overline{LQI} 、 \overline{SNR} 的有效性和组合协方差函数的有效性，与 DBN 模型相比，C-GPR 模型具有更

高的预测准确度。

参考文献:

- [1] 孙佩刚, 赵海, 罗玎玎, 等. 无线传感器网络链路通信质量测量研究[J]. 通信学报, 2007, 28(10): 14-22.
SUN P G, ZHAO H, LUO D D, et al. Study on measurement of link communication quality in wireless sensor networks[J]. Journal on Communications, 2007, 28(10): 14-22.
- [2] BACCOUR N, KOUBÂA A, MOTTOLA L, et al. Radio link quality estimation in wireless sensor networks: a survey[J]. ACM Transactions on Sensor Networks, 2012, 8(4): 1-35.
- [3] PENGWON K, KOMOLMIS T, CHAMPRASERT P. Solving asymmetric link problems in WSNs using site link quality estimators and dual-tree topology[C]//International Conference on Electrical Engineering/electronics, Computer, Telecommunications and Information Technology. 2016: 1-4.
- [4] BAS C U, ERGEN S C. Spatio-temporal characteristics of link quality in wireless sensor networks[C]//IEEE Wireless Communications and Networking Conference. 2012: 1152-1157.
- [5] DEZFOULI B, RADI M, RAZAK S A, et al. Modeling low-power wireless communications[J]. Journal of Network & Computer Applications, 2015, 51(C): 102-126.
- [6] SRINIVASAN K, LEVIS P. RSSI is under appreciated[C]//IEEE the Third Workshop on Embedded Networked Sensors (EmNets). 2006: 239-242.
- [7] TANG L, WANG K C, HUANG Y, et al. Channel characterization and link quality assessment of IEEE 802.15.4-compliant radio for factory environments[J]. IEEE Transactions on Industrial Informatics, 2007, 3(2): 99-110.
- [8] QIN F, DAI X, MITCHELL J E. Effective-SNR estimation for wireless sensor network using Kalman filter[J]. Ad Hoc Networks, 2013, 11(3): 944-958.
- [9] WOO A, CULLER D. Evaluation of efficient link reliability estimators for low-power wireless networks[C]//UCB Technical Report. 2003: 1-20.
- [10] 黄庭培, 李栋, 张招亮, 等. 突发性链路感知的自适应链路质量估计方法[J]. 通信学报, 2012, 33(6): 30-39.
HUANG T P, LI D, ZHANG Z L, et al. Bursty-link-aware adaptive link quality estimation method[J]. Journal on Communications, 2012, 33(6): 30-39.
- [11] COUTO D S J D, AGUAYO D, BICKET J, et al. A high-throughput path metric for multi-hop wireless routing[J]. Wireless Networks, 2005, 11(4): 419-434.
- [12] BOANO C A, ZÚÑIGA M A, VOIGT T, et al. The triangle metric: fast link quality estimation for mobile wireless sensor networks[C]//The 19th International Conference on Computer Communication and Networks. 2010: 1-7.
- [13] BACCOUR N, KOUBÂA A, YOUSSEF H, et al. F-LQE: a fuzzy link quality estimator for wireless sensor networks[C]//European Conference on Wireless Sensor Networks (EWSN). 2010: 240-255.
- [14] LIU T, CERPA A E. Data-driven link quality prediction using link features[J]. ACM Transactions on Sensor Networks, 2014, 10(2): 1-35.
- [15] 舒坚, 汤津, 刘琳岚, 等. 基于模糊支持向量回归机的 WSNs 链路质量预测[J]. 计算机研究与发展, 2015, 52(8): 1842-1851.
SHU J, TANG J, LIU L L, et al. Fuzzy support vector regression-based link quality prediction model for wireless sensor networks[J]. Journal of Computer Research and Development, 2015, 52(8): 1842-1851.
- [16] WENG L N, ZHANG P, FENG Z Y, et al. Short-term link quality prediction using nonparametric time series analysis[J]. Science China Information Sciences, 2015, 58(8): 1-15.
- [17] 舒坚, 刘松, 刘琳岚, 等. 基于动态贝叶斯网络的 WSNs 链路质量预测[J]. 工程科学与技术, 2017, 49(2): 152-159.
SHU J, LIU S, LIU L L, et al. Link quality prediction for WSNs based on dynamic Bayesian networks[J]. Advanced Engineering Sciences, 2017, 49(2): 152-159.
- [18] FEOFUSHING E, KUDELSKI M, NAGI J, et al. Poster abstract: link quality estimation: a case study for on-line supervised learning in wireless sensor networks[C]//The 5th Workshop on Real-World Wireless Sensor Networks. 2014: 97-101.
- [19] 夏战国, 夏士雄, 蔡世玉, 等. 类不平衡的半监督高斯过程分类算法[J]. 通信学报, 2013, 34(5): 42-51.
XIA Z G, XIA S X, CAI S Y, et al. Semi-supervised Gaussian process classification algorithm addressing the class imbalance[J]. Journal on Communications, 2013, 34(5): 42-51.
- [20] CHEN X, WANG H, HUANG J, et al. APU degradation prediction based on EEMD and Gaussian process regression[C]//IEEE International Conference on Sensing, Diagnostics, Prognostics, and Control. 2017: 98-104.
- [21] 孙斌, 姚海涛, 刘婷. 基于高斯过程回归的短期风速预测[J]. 中国电机工程学报, 2012, 32(29): 104-109.
SUN B, YAO H T, LIU T. Short-term wind speed forecasting based on Gaussian process regression model[J]. Proceedings of the Chinese Society for Electrical Engineering, 2012, 32(29): 104-109.
- [22] 李振刚. 基于高斯过程回归的网络流量预测模型[J]. 计算机应用, 2014, 34(5): 1251-1254.
LI Z G. Network traffic forecasting model based on Gaussian process regression[J]. Journal of Computer Applications, 2014, 34(5): 1251-1254.
- [23] 陈艳, 王子健, 赵泽, 等. 传感器网络环境监测时间序列数据的高斯过程建模与多步预测[J]. 通信学报, 2015, 36(10): 252-262.
CHEN Y, WANG Z J, ZHAO Z, et al. Gaussian process modeling and multi-step prediction for time series data in wireless sensor network environmental monitoring[J]. Journal on Communications, 2015, 36(10): 252-262.
- [24] 李军, 张友鹏. 基于高斯过程的混沌时间序列单步与多步预测[J]. 物理学报, 2011, 60(7): 143-152.
LI J, ZHANG Y P. Single-step and multiple-step prediction of chaotic time series using Gaussian process model[J]. Acta Physica Sinica, 2011, 60(7): 143-152.
- [25] 彭虹桥, 顾洁, 胡玉, 等. 基于混沌粒子群—高斯过程回归的饱和和负荷概率预测模型[J]. 电力系统自动化, 2017, 41(21): 25-32, 155.
PENG H Q, GU J, HU Y, et al. Forecasting model of saturated load

based on chaotic particle swarm and optimization-Gaussian process regression[J]. Automation of Electric Power Systems, 2017, 41(21): 25-32,155.

- [26] 乔少杰, 金琨, 韩楠, 等. 一种基于高斯混合模型的轨迹预测算法[J]. 软件学报, 2015, 26(5): 1048-1063.

QIAO S J, JIN K, HAN N, et al. Trajectory prediction algorithm based on Gaussian mixture model[J]. Journal of Software, 2015, 26(5): 1048-1063.

- [27] 耿雪青, 余青山, 韩笑, 等. 基于人工蜂群优化高斯过程的运动想象脑电信号分类[J]. 传感技术学报, 2017, 30(3): 378-384.

GENG X Q, SHE Q S, HAN X, et al. Classification of motor imagery EEG based on Gaussian process optimized with artificial bee colony[J]. Chinese Journal of Sensor and Actuators, 2017, 30(3): 378-384.

- [28] SRINIVASAN K, DUTTA P, TAVAKOLI A, et al. An empirical study of low-power wireless[J]. ACM Transactions on Sensor Networks, 2010, 6(2): 1-49.

- [29] GUNGOR V C, LU B, HANCKE G P. Opportunities and challenges of wireless sensor networks in smart grid[J]. IEEE Transactions on Industrial Electronics, 2010, 57(10): 3557-3564.

- [30] HONG L, LIU X, ZHANG L, et al. Towards sensitive link quality prediction in ad hoc routing protocol based on grey theory[J]. Wireless Networks, 2015, 21(7): 2315-2325.

- [31] 梁宁, 耿立艳, 张占福, 等. 基于 GRA 与 SVM-mixed 的货运量预测方法[J]. 交通运输系统工程与信息, 2016, 16(6): 94-99.

LIANG N, GENG L Y, ZHANG Z F, et al. A prediction method of railway freight volumes using GRA and SVM-mixed[J]. Journal of Transportation Systems Engineering and Information Technology, 2016, 16(6): 94-99.

- [32] 田民, 刘思峰, 卜志坤. 灰色关联度算法模型的研究综述[J]. 统计与决策, 2008, 2008(1): 24-27.

TIAN M, LIU S F, BU Z K. Research summary of grey relational model[J]. Statistics and Decision, 2008, 2008(1): 24-27.

- [33] 万良勇, 刘开云. 公路隧道交通量预测的粒子群高斯过程耦合模型[J]. 北京交通大学学报, 2015, 39(1): 33-39.

WAN L Y, LIU K Y. Traffic flow prediction model of highway tunnel based on PSO-Gaussian process coupled algorithm[J]. Journal of Beijing Jiaotong University, 2015, 39(1): 33-39.

- [34] 崔玮, 吴成东, 张云洲, 等. 基于高斯混合模型的非视距定位算法[J]. 通信学报, 2014, 35(1): 99-106.

CUI W, WU C D, ZHANG Y Z, et al. GMM-based localization algorithm under NLOS conditions[J]. Journal on Communications, 2014, 35(1): 99-106.

[作者简介]



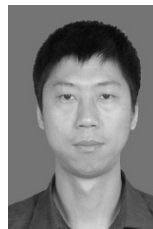
舒坚 (1964-), 男, 江西南昌人, 南昌航空大学教授、硕士生导师, 主要研究方向为无线传感器网络、软件工程等。



刘满兰 (1992-), 女, 湖南耒阳人, 南昌航空大学硕士生, 主要研究方向为无线传感器网络。



尚亚青 (1991-), 女, 河南开封人, 南昌航空大学硕士生, 主要研究方向为无线传感器网络。



陈宇斌 (1977-), 男, 江西南昌人, 南昌航空大学讲师, 主要研究方向为无线传感器网络、高性能计算。



刘琳岚 (1968-), 女, 湖南东安人, 南昌航空大学教授, 主要研究方向为无线传感器网络、分布系统等。