

# 基于随机森林算法 的智能NB-IoT覆盖评估

Intelligent NB-IoT Coverage Evaluation Based on Random Forest Algorithm

曾伟,刘哲,钟检荣(中国联通北京市分公司,北京100038)

Zeng Wei,Liu Zhe,Zhong Jianrong(China Unicom Beijing Branch,Beijing 100038,China)

## 摘要:

介绍了一种基于随机森林算法的智能NB-IoT覆盖评估新方法,通过人工智能算法解决了NB-IoT覆盖评估不精确的难题。首先从理论推导和实测矫正2个方面介绍了目前NB-IoT覆盖评估一般所采用的线性评估方法,然后介绍了基于随机森林算法的数据获取和模型构建,最后对比了NB-IoT覆盖线性评估方法和随机森林算法模型评估方法的结果。

## 关键词:

随机森林算法;窄带物联网;覆盖评估

doi:10.12045/j.issn.1007-3043.2021.08.016

文章编号:1007-3043(2021)08-0080-04

中图分类号:TN919

文献标识码:A

开放科学(资源服务)标识码(OSID):



## Abstract:

A new method of intelligent NB-IoT coverage evaluation based on random forest algorithm is introduced. The problem of inaccurate NB-IoT coverage evaluation is solved by artificial intelligence algorithm. First, it introduces the linear evaluation methods currently used in NB-IoT coverage evaluation from theoretical derivation and measured correction. Then it introduces the data acquisition and model construction of random forest algorithm evaluation methods. Finally, it compares the results of the NB-IoT coverage linear evaluation method and random forest algorithm model evaluation.

## Keywords:

Random forest algorithm;NB-IoT;Coverage evaluation

**引用格式:**曾伟,刘哲,钟检荣. 基于随机森林算法的智能NB-IoT覆盖评估[J]. 邮电设计技术,2021(8):80-83.

## 0 引言

窄带物联网(Narrow Band Internet of Things, NB-IoT)作为5G低功耗大连接场景的前期技术已得到了广泛关注,NB-IoT目前已成为各大运营商物联网采用的主流技术。与传统通信技术相比,窄带物联网技术具有低成本、低功耗的特点,应用范围更加广泛。在技术路径的选择上,窄带物联网技术对传统通信技术做了极大的简化,为了节省终端功耗,不支持终端的

MR测量上报。通过传统手段来获取NB-IoT网络的信号状况,不仅周期长、成本高、效率低,而且部分业务场景点难以覆盖,因此无有效的手段获取网络覆盖状况成为了窄带物联网技术发展中亟待解决的难题。在人工智能日益成熟、应用愈发广泛的今天,在NB-IoT网络覆盖评估引入随机森林算法,有望解决这一难题,为窄带物联网的进一步推广应用扫清障碍。

## 1 NB-IoT覆盖线性评估

在目前的技术手段与实践中除了路测评估之外,利用NB-IoT网络与UMTS900M网络同站址同天馈的

收稿日期:2021-07-03

特点,常采用覆盖线性评估的方法。共天馈小区信号空间传播路径大致相同,路径损耗差基本定于不同频段空间损耗差。基于UMTS900M网络栅格化MR数据构建路损矩阵,并利用发射功率和路径损耗差别,实现对NB-IoT网络栅格的覆盖预测评估。如图1所示。

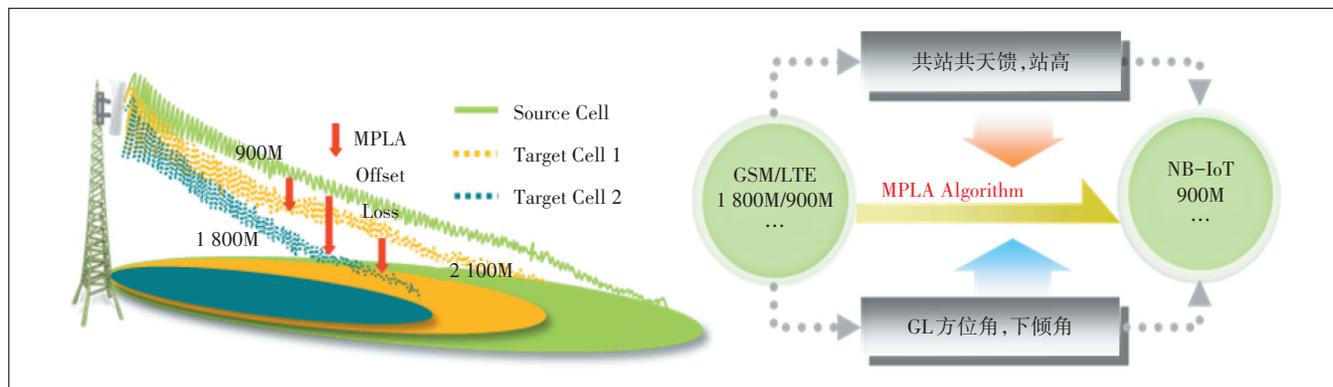


图1 NB-IoT覆盖线性评估原理

接收信号强度差别只与发射功率(Tx Power)和路径损耗(PathLoss)有关。发射功率差异如表1所示。

表1 发射功率差异

制式	总功率	参考信号/导频功率
NB-IoT	2×10 W(1个RB)	32.2 dBm
UMTS 900	20 W(导频功率为总功率10%)	33 dBm

路径损耗可以由自由空间传播损耗公式计算得到。自由空间传播损耗公式:

$$Los = 32.44 + 20\log d(\text{km}) + 20\log f(\text{MHz}) \quad (1)$$

由式(1)可知,在传播距离相同的情况下,路径损耗只与工作频率相关,中国联通NB-IoT网络使用的频段为900 MHz与U900网络同频段,所以2张网络的路径损耗一致。

综上所述,如用U900网络作为NB-IoT网络覆盖线性评估的参照物,那么在同站址共天馈场景下可计算N900网络覆盖强度:

$$N900 \text{ 接收信号强度} - U900 \text{ 接收信号强度} = 32.2 - 33 = -0.8 \text{ dB}; \text{即 } N900 \text{ RSRP} = U900 \text{ RSCP} + 0.8 \text{ dB}$$

## 1.2 实测矫正

1.1节通过理论说明了通过U900的覆盖强度来估算NB-IoT网络覆盖强度的可行性,U900的覆盖电平与N900的覆盖电平存在规律差异。在实际应用中,

## 1.1 理论推导

由于NB-IoT网络与UMTS网络采用1:1同站址组网,共用天馈系统,2张网络在射频特性上,如天线挂高、方位角、天线角等保持一致。所以,NB-IoT网络与UMTS网络的天馈损耗(Feeder loss)和天线增益(An-

由于网管平台采集及解析U900 MR数据产生的偏滞,U900与N900的覆盖差值还需要经过实测矫正。选取了一片区域进行验证,对45个测试点的U900与N900进行CQT测试,对比测试结果,如图2所示。

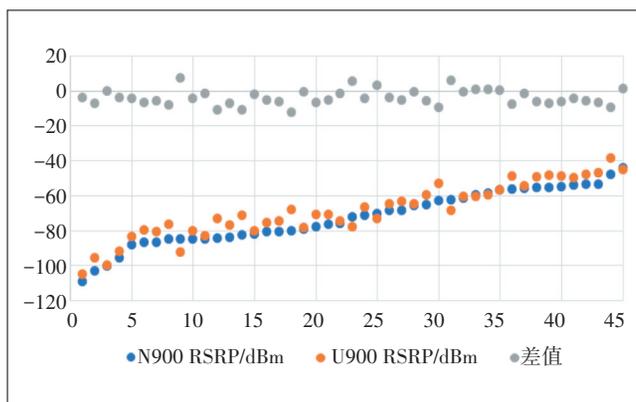


图2 N900 RSRP与U900 RSCP的实测差值示意图

从图2可以看出,U900覆盖电平值与N900覆盖电平值整体差值保持平稳,确实存在规律性差异,差异值平均为4 dB,即U900 RSCP比N900 RSRP高约4个dB,因此N900 RSRP≈U900 RSCP + 4 dB。

## 2 随机森林模型的建立

随机森林是一种统计学习理论,是利用bootstrap重抽样方法从原始样本中抽取多个样本,对每个bootstrap样本进行决策树建模,然后组合多棵决策树的

结果,通过投票得出最终结果。随机森林模型具有很高的预测准确率,对异常值和噪声具有很好的容忍度,且不容易出现过拟合。

### 2.1 数据的来源与选取

本文以目标点位处的U900制式的RSCP值、U900制式的EcIo值、L1800制式的RSRP值、L1800制式的SINR值、U2100制式的RSCP值、U2100制式的EcIo值、经度、纬度及采样点与主服务基站距离这9个指标作为输入特征;目标点位的NB-IoT覆盖电平值作为输出特征。

在原始样本集中,输入特征数据可以直接从网管后台提取相关指标,输出特征通过实地测试的方法获得。同样的,在评估待测点位的NB-IoT覆盖电平值时,只需要从后台网管提取待测点位的9个输入特征值即可。

### 2.2 随机森林算法模型

随机森林算法是一种集成学习算法,综合多个决策树的预测结果得到最终结果。随机森林中随机是核心,通过随机的选择样本、特征,降低了决策树之间的相关性。随机森林中的随机主要有2层意思,一是随机在原始训练数据中有放回地选取等量的数据作为训练样本,二是在建立决策树时,随机地从特征中选取一部分特征建立决策树。这2种随机使得各个决策树之间的相关性较小,可进一步提高模型的准确性。随机森林算法分为如下3个步骤。

步骤1,生成原始数据样本集,模型输入量的个数为9个,输出量的个数为1个。选取 $N$ 个NB-IoT网络终端点位来生成原始数据样本集,一条原始数据样本包含9个输入特征值和1个输出值。原始数据样本集包含 $N$ 条原始数据样本。

步骤2,用步骤1生成的原始数据样本集通过有放回抽样 $M$ 次生成训练样本集,建立窄带物联网覆盖评估模型,评估模型以随机森林算法为核心,由 $M$ 棵决策树组成。每棵决策树都由一个训练样本集生成,决策树生成时从根节点开始,采用最小二乘偏差(LSD)方法递归确定每个节点的分裂变量和分裂点,直到满足停止条件,如图3所示。

步骤3,获取待测点位的9个模型输入数据,通过窄带物联网覆盖评估模型计算出待测点位的信噪比值。将9个模型输入数据输入模型,每棵决策树都可以得出1个信噪比值,记为 $f_1(x), f_2(x) \dots f_m(x)$ , $M$ 棵决策树所得结果的平均值即为模型的输出结果。即:

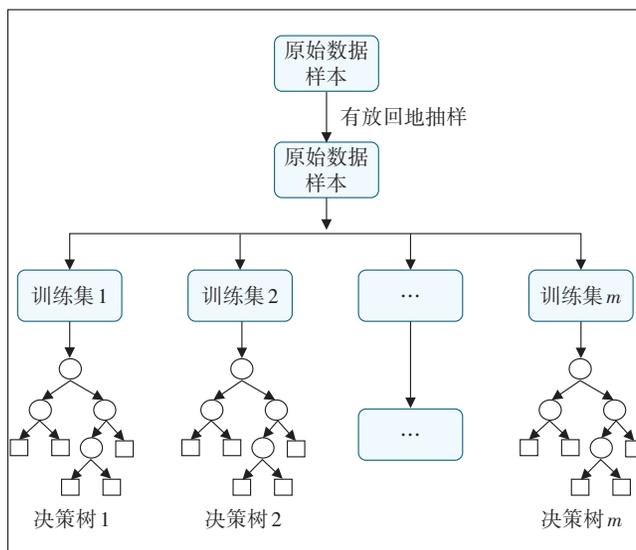


图3 随机森林模型生成图

$$f(x) = \sum_{i=1}^m f_i(x) \quad (2)$$

在步骤2中,训练样本集的生成方法是:从 $N$ 条原始样本数据集中,有放回随机抽样 $n(n \leq N)$ 条原始数据样本,再从每条原始数据样本中随机选取 $k(k \leq 9)$ 个输入特征值与1个输出值,生成1个训练样本集。1个训练样本集为一个 $n \times (k + 1)$ 的数组。

生成决策树时的最小二乘偏差(LSD)方法是求解:

$$\min_{t^i, t_j^i} \left[ \min_{c_1} \sum_{x_i \in R_1(t^i, t_j^i)} (y_i - c_1)^2 + \min_{c_2} \sum_{x_i \in R_2(t^i, t_j^i)} (y_i - c_2)^2 \right] \quad (3)$$

其中, $t^i$ 是分裂特征, $t_j^i$ 是分裂点, $R_1, R_2$ 是按照分裂点 $t_j^i$ 划分的2个区域:

$$R_1(t^i, t_j^i) = \{x | x^{(t^i)} \leq t_j^i\}, R_2(t^i, t_j^i) = \{x | x^{(t^i)} > t_j^i\} \quad (4)$$

$c_1, c_2$ 是 $R_1, R_2$ 两区域内的平均值:

$$c_1 = \text{avg}(y_i | x_i \in R_1(t^i, t_j^i)), c_2 = \text{avg}(y_i | x_i \in R_2(t^i, t_j^i)) \quad (5)$$

$x_i, y_i$ 是模型的输入和输出, $i \in \{1, 2, \dots, n\}$ ,  
 $j \in \{1, 2, \dots, k\}$ 。

遍历特征 $t^i$ ,扫描分裂点 $t_j^i$ ,选择使最小二乘偏差达到最小值的分裂特征和分裂点 $(t^i, t_j^i)$ ,作为决策树的分裂依据。

为了控制过拟合现象的发生,可以把决策树生成停止条件指定为:某一分裂点左右任意一支的数据集样本数小于 $5\% \times n$ 。

### 3 结果对比分析

本文构建了2种NB-IoT覆盖评估模型,选取同一区域对线性评估模型和随机森林模型的效果进行验证,通过评估值与实测值的对比,来验证模型的准确性,比较模型的性能。如图4和图5所示。

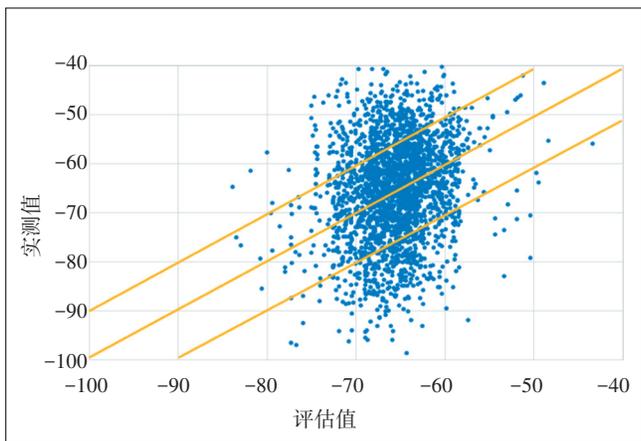


图4 线性评估模型NB-IoT覆盖评估值与实测值对比图

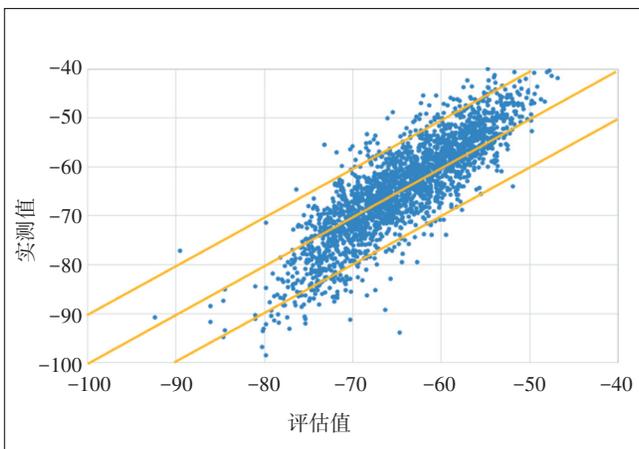


图5 随机森林模型NB-IoT覆盖评估值和实测值对比

结果表明,由线性评估模型推算出的评估值与实测值大部分偏差在20 dB以内,标准偏差 $\sigma = 10.5$ ,评估结果较为一般。而由随机森林模型推算得出的评估值与实测值有显著的相关性,标准偏差 $\sigma = 5.7$ ,在覆盖电平值较差的区域采样点偏差大,这与NB-IoT弱覆盖区域采样点较少有关,如能增加采样点补充进训练样本集,则评估精度还可以进一步提升。

综上所述,由随机森林模型得到的评估值较为理想,弥补了NB-IoT网络覆盖评估不准确的漏洞,具有市场推广意义。

### 4 结束语

本文深入研究NB-IoT覆盖的评估方法,结合人工智能领域的随机森林算法,探索出一种开展NB-IoT覆盖评估的有效方法。对比线性差值评估方法,有效提高了覆盖评估精度,大大提高了覆盖预测评估效率。

对NB-IoT物联网覆盖的精准评估,可以有效地指导NB-IoT网络规划建设和优化运维,精准助力业务部门对潜在用户的售前牵引,为运营商物联网业务的部署应用提供了有力支撑。NB-IoT点位的评估优化,支撑了包括北京联通亦庄智能井盖项目、丰台水位监控项目以及新冠疫情期间隔离报警设备项目在内的一批重点项目,均取得了显著的效果。

#### 参考文献:

- [1] 肖亚,宋知明.基于MPLA算法的NB-IoT网络覆盖评估[J]. 通信技术,2018(7):9-12.
- [2] 王文哲,解解等.基于创新MR技术的NB-IoT覆盖评估新体系[J]. 邮电设计技术,2019(1):31-36.
- [3] 勾保同,赵建平.NB-IOT的覆盖增强技术探讨[J]. 通信技术,2018(6):1254-1258.
- [4] 黄润.NB-IoT低速率窄带物联网通信技术现状及发展趋势[J]. 计算机产品与流通,2017(8):58-58.
- [5] 王磊,宋锴.NB-IoT规划部署策略研究[J]. 电信工程技术与标准化,2018(11):72-76.
- [6] 夏学成.NB-IOT发展概述[J]. 广东通信技术,2017(6):6-9.
- [7] 林芳.NB-IoT物联网覆盖增强技术及在远程抄表系统中的应用[J]. 电子世界,2017(14):124-124.
- [8] 黄悦,汤远方.NB-IoT物联网组网及覆盖能力探讨[J]. 移动通信,2017(18):11-15.
- [9] 方匡南,吴见彬.随机森林方法研究综述[J]. 统计与信息论坛,2011(3):32-38.
- [10] 姚登举,杨静.基于随机森林的特征选择算法[J]. 吉林大学学报:工学版,2014(11):137-141.
- [11] 王丽爱,马昌.基于随机森林回归算法的小麦叶片SPAD值遥感估算[J]. 农业机械学报,2015(1):259-265.
- [12] 范昕,赵桂新.使用随机森林判别分析法预测黑加仑油胶囊治疗高血脂的效果[J]. 中医药信息,2012(4):43-46.
- [13] 王丽婷,丁晓青.基于随机森林的人脸关键点精确定位方法[J]. 清华大学学报:自然科学版,2009(4):543-546.

#### 作者简介:

曾伟,高级工程师,硕士,主要从事移动通信网络规划及优化工作;刘哲,工程师,硕士,主要从事移动通信网络优化工作;钟检荣,高级工程师,硕士,主要从事移动通信网络优化工作。