

基于K最近邻算法的地铁4G 室内故障定位方法研究

Research on Fault Locating Method of Metro 4G Indoor Distribution System
Based on K Nearest Neighbor Algorithm

滕祖伟,周杰华,肖 波(中国联通湖北分公司,湖北 武汉 430040)
Teng Zuwei,Zhou Jiehua,Xiao Bo(China Unicom Hubei Branch,Wuhan 430040,China)

摘 要:

针对电信运营商无法直接对空间封闭的地铁隧道进行栅格化操作的难点,创新性采用空间相关性测量方法,对地铁隧道进行相同时间间隔的连续分割并建立位置指纹库。然后,采用K最近邻算法把位置指纹库与实时上报的MR数据进行匹配,一旦连续多个地理栅格都未匹配到MR数据,则判定该区域室内系统存在故障。该定位方法实现简单,定位精度高。

关键词:

K最近邻算法;地铁;室内系统;故障定位
doi: 10.12045/j.issn.1007-3043.2019.06.012
中图分类号:TN929.5
文献标识码:A
文章编号:1007-3043(2019)06-0052-04

Abstract:

In view of the difficulty of the grid operation of the metro tunnel which is closed to space directly by the telecom operators, the spatial correlation measurement is innovatively adopted to complete the continuous segmentation of the metro tunnel at the same time interval and establish the location fingerprint library. Then, the K nearest neighbor algorithm is used to match the position fingerprint library with the real-time MR data, and once the continuous multiple geographic grids are not matched to the MR data, the fault of indoor distribution system is determined. The positioning method is simple, and the positioning accuracy is high.

Keywords:

K nearest neighbor algorithm; Metro; Indoor distribution system; Fault locating

引用格式:滕祖伟,周杰华,肖波. 基于K最近邻算法的地铁4G室内故障定位方法研究[J]. 邮电设计技术,2019(6):52-55.

0 引言

地铁已经成为城市最常用的公共出行方式,4G用户经常使用移动互联网来消磨时间,因而保障地铁4G用户感知愈发重要。2017年中国联通集团要求:在保持重点城市网络质量的基础上,打造地铁等4G室内网络领先优势。目前地铁室内分布系统主要由有源设备及无源分布系统组成,其中无源分布系统由馈线、无源器件、天线组成来实现射频信号在建筑物内的传输及覆盖。后台告警仅能监控室内信源设备故障,无

法监控到分布系统等无源器件问题,这些无源器件问题主要依靠巡检和投诉来定位,不仅周期长,且故障定位繁琐、工作量大、定位准确率低。

K最近邻算法是一个理论上比较成熟的机器学习方法,其核心思想是:给定一个训练样本集,输入没有标签的新数据,与样本集中数据对应的特征进行比较,然后选择样本集中距离最近邻的 k 个分类标签,并按照一定规则确定最终的分类标签。通常采用最小欧氏距离或最大余弦距离衡量。

MR是手机在执行业务过程中向网络上报的、反映用户当前无线环境的下行信号强度和品质等信息的报告,其上报周期一般设定为5 s。

收稿日期:2019-04-30

现有的基于K最近邻算法和位置指纹库的MR定位技术已广泛应用于移动无线网质量评估、网络规划、网络优化工作中。但地铁的室内场景比较特殊,不仅其中的终端经纬度信息无法获取,而且窄长型的隧道环境封闭,作为第三方的电信运营商无法直接对其空间进行栅格化操作,因此,无法有效实现对地铁室分系统故障的精准定位。

1 主要思路

针对电信运营商无法直接对空间封闭的地铁隧道进行栅格化操作的难点,创新性采用空间相关性测量方法,对地铁隧道进行相同时间间隔的连续分割,完成地理栅格化并建立位置指纹库。然后,采用K最近邻机器学习算法,把位置指纹库与实时上报的MR数据进行匹配:一旦连续多个地理栅格都未匹配到MR数据,则认为该区域室分系统存在故障。

其中,地铁线路内有多列列车连续、匀速运行,相邻列车时间间隔较短(通常低于5 min),可以认为:地铁手机用户持续占用整条地铁的室分系统,那么每次周期上报的MR数据就是针对整条地铁室分系统无线信号进行的一次“快照”。同时,由于地铁室分系统信号与室外信号隔离好,具有干扰小、信号电平波动小的特点,因此K最近邻算法定位准确度高。

2 定位过程

2.1 流程图

故障定位流程图如图1所示。

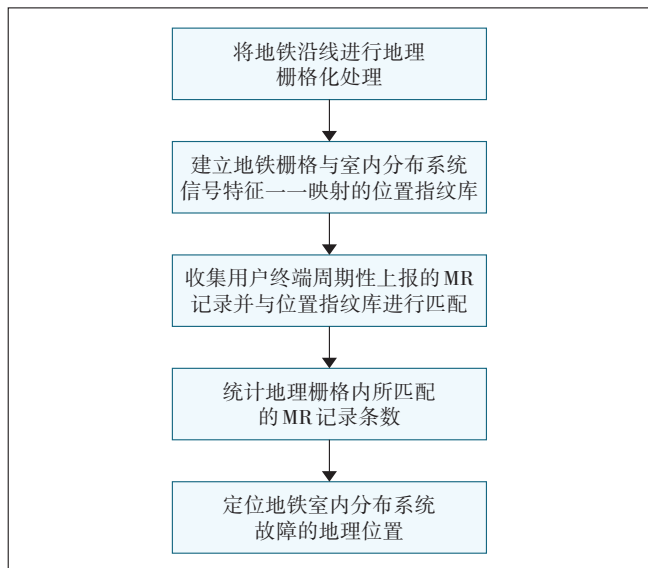


图1 故障定位流程图

2.2 具体实施

武汉地铁3号线全线位于地下,起于宏图大道,止于沌阳大道,全长29 640 m,沿线24个地铁站。覆盖该地铁沿线共使用4G基站29个,4G小区65个。下面以武汉地铁3号线为例进行具体实施说明。

2.2.1 将地铁沿线进行地理栅格化处理

a) 执行空间相关性测量生成CQT文件:测试人员携带已安装前台测试软件Pilot Walktour的测试终端iPhone 6S A和测试终端iPhone 6S B,从武汉地铁3号线起点宏图大道站台进入地铁列车尾部同时开始CQT测试。当列车启动、离开站台或到达下一站台停车时,测试人员要在已导入测试终端A的地铁站台分布图上进行手动打点(注:每次手动打点不要重叠在一起),一直持续到地铁终点沌阳大道站台,从而得到测试终端A、B生成的2个rcu格式CQT文件(见图2)。

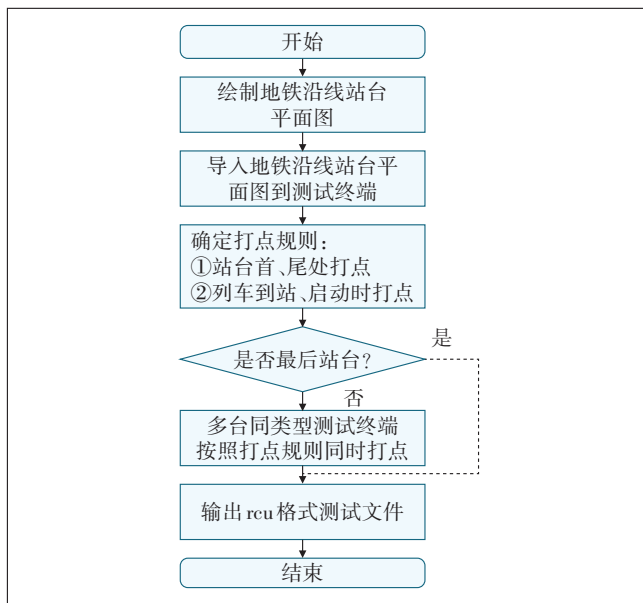


图2 空间相关性测量过程流程图

b) 对地铁沿线进行相同时间间隔的连续分割:测试人员采用Pilot Walktour软件解析测试终端A生成的CQT文件,并输出TXT文本格式的地铁室内分布系统信号特征记录。然后,根据之前列车停车、启动时的手动打点,剔除列车停靠站台时的信号特征记录,完成对整条地铁200 ms间隔周期(即测试终端A的芯片采样周期)的均匀等分,并得到每个200 ms所对应的信号特征记录信息:采样时刻、主小区唯一标识、主小区信号电平、邻小区PCI、邻小区信号电平。

c) 针对这些信号特征记录从“1”开始进行顺序编号,得到长度为29 640 m的武汉地铁3号线全部11 694

个地理化栅格(见图3)。



图3 地铁3号线地理栅格化

d) 采用上述相同步骤处理测试终端B生成的CQT文件,实现测试终端B对整条地铁站台和隧道的地理栅格化处理。

2.2.2 建立地理栅格与室内分布系统信号特征一一映射的位置指纹库

针对测试终端A、测试终端B产生的相同栅格编号所对应的主小区信号电平、邻小区信号电平进行算术平均,得到对应栅格内的信号电平强度向量(包含主小区信号电平强度平均值、邻小区信号电平强度平均值),从而建立地铁地理栅格与室内分布系统信号特征一一映射的位置指纹库(见表1)。

2.2.3 采用K最近邻算法实时定位MR记录

表1 位置指纹库

栅格编号	信号电平个数	信号电平序号	主小区唯一标识	PCI	信号电平/dBm
1	3	1	HUB_HW_OMC1.107191.26	421	-84.56
1	3	2		307	-96.56
1	3	3		394	-101.18
2	3	1	HUB_HW_OMC1.107191.26	421	-83.44
2	3	2		307	-96.56
2	3	3		394	-100.87
3	2	1	HUB_HW_OMC1.107191.26	421	-84.13
...
11 691	4	1	HUB_HW_OMC1.107164.27	386	-67.31
11 692	4	2		385	-79.43
11 693	4	3		335	-81.37
11694	4	4		392	-100.23

K最近邻算法实时定位MR记录如图4所示。

采集并解析地铁室内分布系统OMC周期生成的15 min粒度的MR原始文件,得到地铁手机用户实时上报的待匹配MR记录中的主小区唯一标识、主小区信号电平、邻小区PCI及邻小区信号电平,并从位置指纹库中筛选出具有相同主小区唯一标识和邻小区PCI的1个或多个信号电平强度向量。

然后,采用K最近邻算法计算这些信号电平强度向量与待匹配MR的信号电平强度向量之间的欧氏距离。根据计算得到的欧氏距离和距离门限阈值确定最相似的k个栅格,并按照一定规则从k个栅格中选择1个栅格作为待匹配MR记录的栅格位置估计值(见表2)。

2.2.4 统计地理栅格内所匹配的MR记录条数

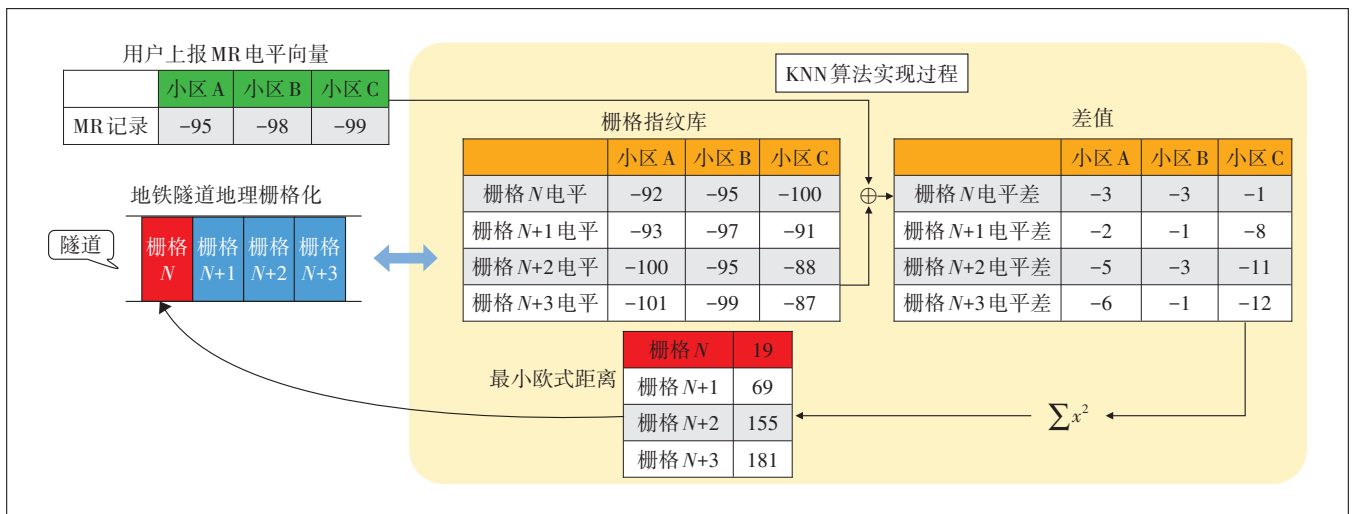


图4 K最近邻算法实时定位MR记录

表2 栅格位置估计值

栅格估计值	MR上报时间	MR上报主小区唯一标识	MR上报信号电平/dBm
3	T11:53:53.920	HUB_HW_OMC1.107191.26	-80
7	T11:39:44.000	HUB_HW_OMC1.107191.26	-74
7	T11:41:05.920	HUB_HW_OMC1.107191.26	-68
9	T11:46:43.840	HUB_HW_OMC1.107191.26	-76
11	T11:31:32.480	HUB_HW_OMC1.107191.26	-77
11	T11:47:45.280	HUB_HW_OMC1.107191.26	-65
11	T11:52:32.000	HUB_HW_OMC1.107191.26	-74
...
11 680	T11:50:29.120	HUB_HW_OMC1.107164.28	-80
11 684	T11:11:24.160	HUB_HW_OMC1.107164.27	-83
11 687	T11:24:42.880	HUB_HW_OMC1.107164.27	-80
11 692	T11:35:28.000	HUB_HW_OMC1.107164.27	-77

按小时粒度统计武汉地铁3号线全部11 694个地理栅格内所匹配的MR记录条数(见表3)。

表3 栅格所匹配的MR记录条数

栅格编号	日期	时间点	匹配的MR记录条数
4	2018/4/2	10	3
5	2018/4/2	10	10
7	2018/4/2	10	9
11	2018/4/2	10	2
12	2018/4/2	10	31
13	2018/4/2	10	2
14	2018/4/2	10	17
...
11 675	2018/4/2	10	2
11 678	2018/4/2	10	5
11 680	2018/4/2	10	12
11 687	2018/4/2	10	6

2.2.5 定位地铁室内分布系统故障的地理位置

如果一个小时内未匹配到MR记录的地理栅格序号连续,并且序号连续的栅格数量大于设定门限值(50)时,可判定这些连续地理栅格所对应的室内分布系统覆盖区域出现故障。

2.3 效果验证

随机选取武汉地铁3号线的宏图大道站、市民之家站2个站台进行定位精准性测试。

每个站台起始位置测试1 min,间隔10 m测试1 min,依次重复10次,记录10组的时间点和位置点,共进行20次测试。同时后台跟踪测试手机的MME UE S1AP ID,从而可找到对应MR记录并匹配到栅格

位置。最后,统计栅格位置和实际位置的距离偏差:定位偏差在-20~+20 m的MR数量占比为86.34%(见表4)。

表4 MR定位精度统计

位置序号	上报MR数量	定位偏差在-20~+20 m的MR占比/%	位置序号	上报MR数量	定位偏差在-20~+20 m的MR占比/%
1	12	91.67	11	12	91.67
2	12	83.33	12	13	66.67
3	13	92.31	13	10	100.00
4	11	100.00	14	13	92.31
5	12	83.33	15	12	91.67
6	12	90.91	16	12	83.33
7	12	91.67	17	12	91.67
8	11	83.33	18	12	92.31
9	13	50.00	19	12	66.67
10	12	92.31	20	12	91.67
偏差在-20~+20 m的MR数量占比/%					86.34

2.4 应用情况

统计某时段武汉地铁3号线每个地理栅格内所匹配的MR条数,发现4处疑似故障。其中,宏图大道站→市民中心站之间的隧道内86个栅格(约250 m)未匹配到MR数据,经核实:信源正常,但POI存在上行硬件故障。协调铁塔替换后,恢复正常。

3 结论

本文描述的地铁室内分布系统故障的定位方法,实际上是位置指纹库定位算法的逆向应用,可拓展应用到地停、高铁隧道、高速隧道等其他密闭场景的故障定位。如果开启中国联通4G基站的异频测量功能,可得到密闭场景中其他运营商(中国移动、中国电信)的信号情况,找到相比竞争对手的优势和劣势,有的放矢地指导日常优化工作。

参考文献:

- [1] 李航. 统计学习方法[M]. 1版. 北京:清华大学出版社,2012:38-40.
- [2] 周志华. 机器学习[M]. 北京:清华大学出版社,2016:225-226.
- [3] 胡应添,李伟丹,郝秀娟. 新型室分系统定位方法的研究[J]. 移动通信,2016(14):56-59.

作者简介:

滕祖伟,硕士,主要研究方向为网络优化、数据挖掘;周杰华,学士,主要研究方向为网络优化、特殊场景移动网络信号规划及优化;肖波,学士,主要研究方向为网络优化、移动网络信号规划及优化。