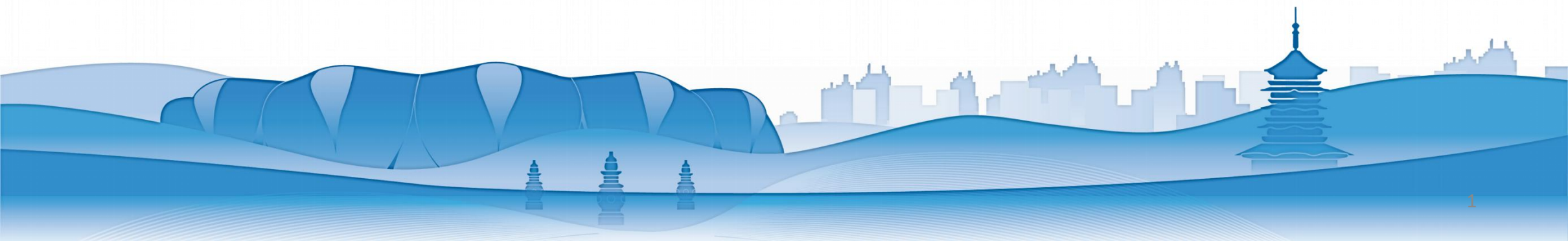


# 5G<sup>++</sup>创新实训基地

技术实操的练兵场 · 能力认证的人才站 · 5G应用的孵化器





中国移动  
China Mobile

5G<sup>+</sup>创新实训基地  
技术实操的练兵场 · 能力认证的人才站 · 5G应用的孵化器

## 版 权 声 明

本课程系由中国移动通信集团浙江有限公司（简称“浙江移动”）受中国移动通信集团有限公司委托开发，版权归属浙江移动，并受法律保护。转载、摘编或利用其它方式使用本课程文字或者观点的，应注明“来源：中国移动通信集团浙江有限公司”。违反上述声明者，浙江移动将追究其相关法律责任。



中国移动  
China Mobile

5G<sup>+</sup>创新实训基地  
技术实操的练兵场 · 能力认证的人才站 · 5G应用的孵化器

# 基于小区画像的数据挖掘建模

2020年11月

诺基亚-廖文哲

1

课题背景

2

数据获取

3

数据处理与算法设计

4

算法评价与结果呈现

目前，随着网络的覆盖范围越来越广，其覆盖也会越来越深，而随着移动通信技术的发展，其频谱利用带宽也会加大，所以随之所带来的干扰也会越来越多。网络覆盖的模型和层次的复杂度，与干扰排查和处理的难度是成正比的。本提案主要针对以上缺点，使用大数据及机器学习技术，提前主动预警干扰区域，同时依据干扰特征，提供相关的干扰排查思路和步骤，减少干扰排查工作的困难。

干扰类型	干扰产生原因
GPS失锁干扰	GPS失锁、星卡故障、GPS天线故障等原因导致时钟不同步的A基站发射信号干扰到了B基站的上行接收。
超远同频干扰	远距离的站点信号经过传播，DwPTS与被干扰站的UpPTS对齐，导致干扰站的基站发对被干扰站的基站收的干扰。
帧失步干扰	帧偏置配置不当、子帧配比不一致等原因会导致基站间的上下行帧对不齐，导致SiteA的下行干扰到了SiteB的上行，形成帧失步干扰。
重叠覆盖干扰	A小区和B小区存在重叠区域（同频邻区必然会存在一定的切换区域），由于两个小区之间的信号不是一致的，不正交，会形成干扰。
硬件故障干扰	设备故障是指在设备运行中，设备本身性能下降等造成干扰包括：RRU故障，RRU接收链路电路工作异常产生干扰；天馈系统故障，包括天线通道故障，天线通道RSSI接收异常等，天馈避雷器老化，质量问题，产生互调信号落入工作带宽内。

干扰类型	干扰产生原因
杂散干扰	由于发射机中产生辐射信号分量落入受害系统接收频段内，导致受害接收机的底噪抬升，造成灵敏度损失，称之为杂散干扰。 干扰来源：1800MHzWLAN、DSC1800等等。
互调/谐波干扰	不同频率的发射信号形成互调/谐波产物。这些产物落于受害系统接收机频段内或频段相近导致干扰，称之为互调谐波干扰。 干扰来源：GSM900高话务小区；
阻塞干扰	由于强度较大的干扰信号在接收机的相邻频段注入，使受害接收机链路的非线性器件产生失真，甚至饱和，称之为阻塞干扰。 干扰来源：电信FDD-LTE阻塞、学校放置的干扰器等等。
MMDS干扰	MMDS在波段2520—2600MHz且连续发射，而现D频段为2575-2635MHz，当MMDS使用频段在D频段范围内，现网D频段设备则会受到干扰。
伪基站干扰&屏蔽器干扰	伪基站独立于运营商的网络，无法准确同步，伪基站功率很大，又和运营商的网络同频率，导致运营商正常的广播被干扰；屏蔽器在工作过程中以一定的速度从前向信道的低端频率向高端扫描。使手机接收报文信号中形成乱码干扰，手机不能检测出从基站发出的正常数据，使手机不能与基站建立联接。



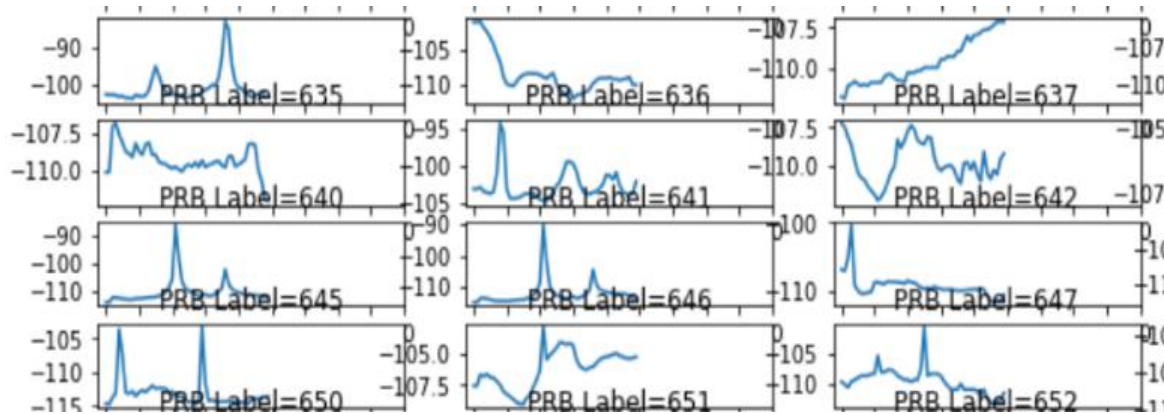
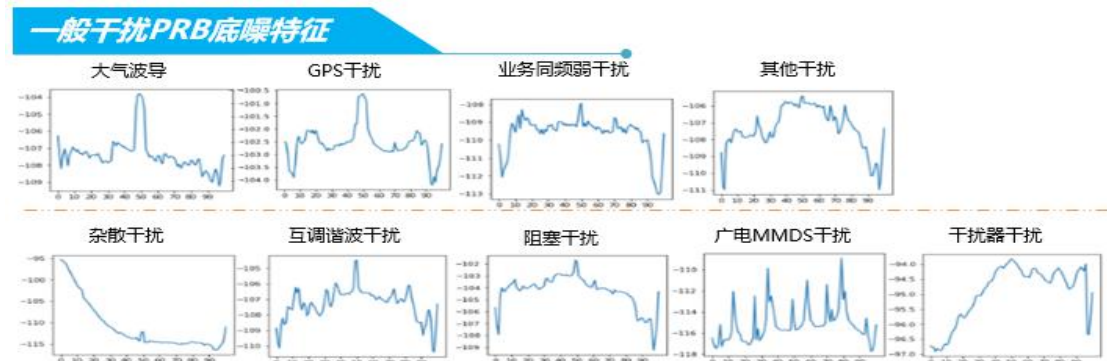
Step1: 通过手工标记已知干扰小区的波形作为标记数据；通过性能中心数据采集15忙时100PRB平均值 $\geq -110$  &  $\leq -80$ 的干扰小区全天100PRB电平数据作为未标记数据，并对缺失值做均值填充。

	CLASS_FRC	CLASS_BIG	CLASS_DE	BANDWID	CLASS	CGI	ULMEANNL_PRBO
	华为第一批	0	0	10	其它干扰	460-00-1023761-134	-111.4387755
	地市9月反	0	1	10	其它干扰	460-00-1023902-129	-115.0645353
	地市9月反	0	2	10	干扰器干扰	460-00-1027290-132	-109.6422354
	地市9月反	0	3	10	干扰器干扰	460-00-1035020-134	-118.7275996
	地市9月反	0	4	10	干扰器干扰	460-00-1035082-132	-115.3895737
	地市9月反	0	5	10	干扰器干扰	460-00-1035258-132	-111.4990209
	地市9月反	0	6	10	干扰器干扰	460-00-1035444-132	-114.3426341
	地市9月反	0	7	10	干扰器干扰	460-00-1035445-132	-118.8910398
	地市9月反	0	8	10	干扰器干扰	460-00-1035447-132	-116.1286085
	地市9月反	0	9	10	干扰器干扰	460-00-197000-193	-74.53367
	华为第一批	0	10	10	其它干扰	460-00-313615-130	-96.1246142
	地市9月反	0	11	10	干扰器干扰	460-00-320371-133	-110.9963923
	地市9月反	0	12	10	干扰器干扰	460-00-320632-163	-115.2425332
	地市9月反	0	13	10	干扰器干扰	460-00-320632-164	-85.6509887
	地市9月反	0	14	10	干扰器干扰	460-00-327322-134	-105.8032857
	华为第三批	0	15	10	大气波导	460-00-395473-10	-71.75
	地市9月反	0	16	10	干扰器干扰	460-00-466949-193	-117.0573127

标记小区数据

小区CGL	上行平均干扰电平	PRBO_上行	PRB1_上行	PRB2_上行	PRB3_上行	PRB4_上行	PRB5_上行
460-00-325567-	-107.365628	-112.802	-115.229	-115.792	-110.396	-110.385	-98.6042
460-00-405857-	-119.919169	-119.063	-120	-119.99	-119.323	-119.094	-119.99
460-00-600181-	-117.806774	-116.51	-117.531	-117.948	-116.979	-117.844	-118
460-00-769770-	-117.268019	-115.24	-117.156	-117.417	-116.406	-116.885	-117.083
460-00-600206-	-107.466564	-107.802	-108.375	-110.802	-109.354	-108.188	-108.01
460-00-600206-	-109.637394	-108.604	-111.792	-110.813	-109.573	-109.719	-109.156
460-00-731069-	-116.847392	-115.302	-115.135	-117.427	-116.51	-115.615	-115.646
460-00-406408-	-116.057607	-116.365	-117.875	-117.948	-107.594	-107.292	-114.854
460-00-406408-	-117.128541	-117.24	-117.938	-118.052	-109.938	-109.698	-117.333
460-00-405937-	-118.224477	-117.688	-118.823	-118.604	-117.604	-118.396	-118.875
460-00-405937-	-116.081043	-115.563	-116.917	-117.125	-114.604	-115.76	-116.563
460-00-752595-	-114.87427	-111.063	-115.042	-115.49	-113.219	-113.375	-113.781
460-00-406719-	-119.709899	-120	-120	-120	-119.583	-119.698	-119.792
460-00-406770-	-115.722813	-113.104	-115.333	-114.417	-114.042	-115.583	-116.698
460-00-600683-	-107.882498	-108.688	-116.26	-117.833	-106.531	-106.781	-106.844
460-00-730873-	-112.226976	-105.708	-110.854	-111.698	-103.948	-103.365	-103.615
460-00-730879-	-113.675521	-112.406	-116.281	-115.354	-111.615	-111.948	-111.771

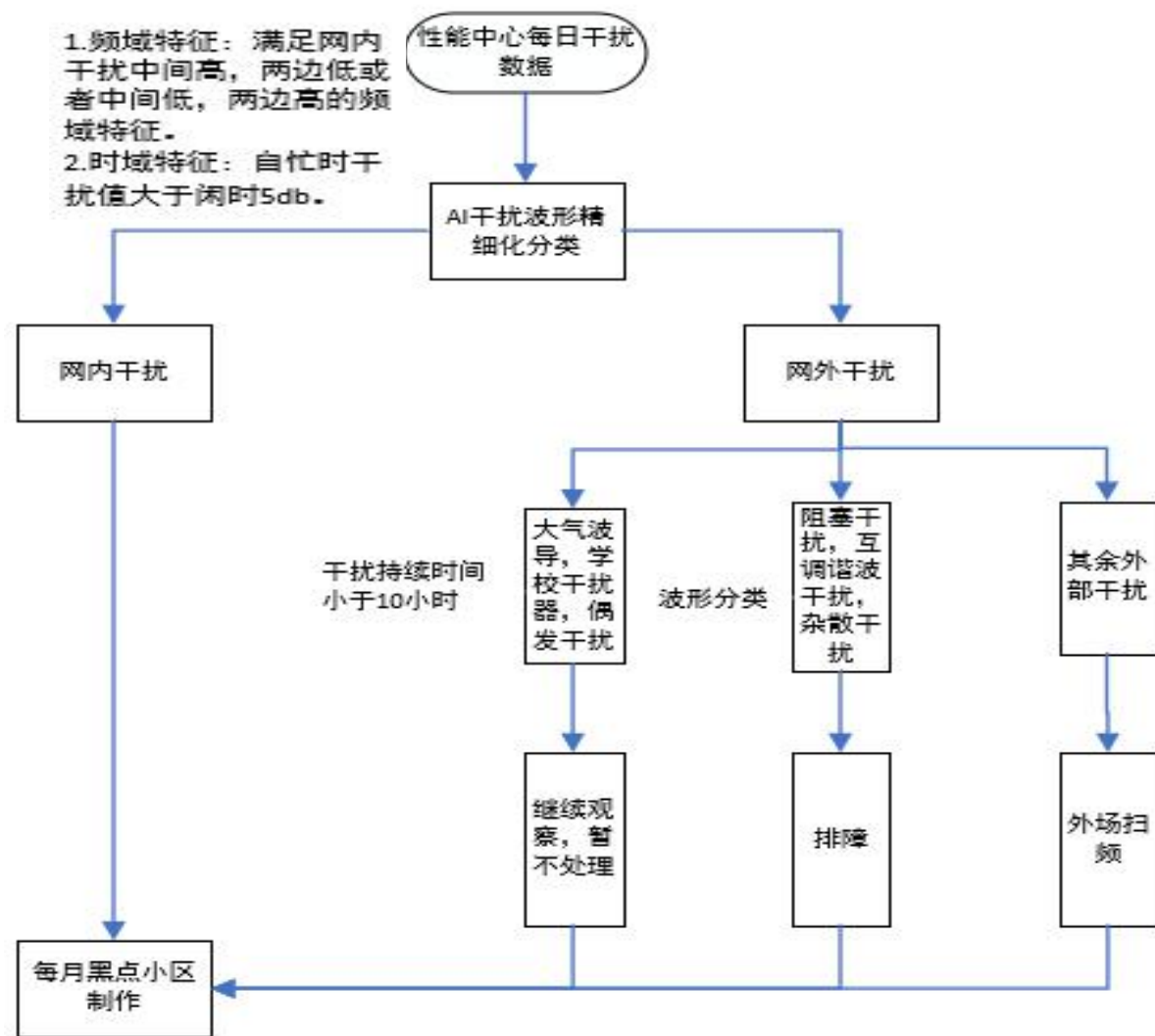
未标记小区数据



**Step2:** 对Step1得到的数据中的缺失值采用均值方式进行处理，即获取小区一天内的干扰数据，取其数据分布的均值填补缺失数据。然后将每个已标记干扰小区级波形作为一个单一类别，利用SEMI-Kmeans算法干扰数据的精细化干扰分类。

**Step3:** 由于网内干扰的特性：**1.频域特征：**满足网内干扰中间高，两边低或者中间低，两边高的频域特征。**2.时域特征：**自忙时干扰值大于闲时5db。利用Step2处理过后的数据，筛选出满足频域特征的波形，并加入时域特征筛查，最终筛选出网内干扰小区。

**Step4:** 利用Step2和Step3筛选处理过后的数据，选出Step2中标记为杂散干扰，阻塞干扰与互调谐波的小区，作为后续排障使用。在剩下的小区中将一天内干扰小时数<10的小区标为大气波导干扰，偶发干扰，学校干扰器。然后将剩下的干扰小区标记为其他干扰





1.轮廓系数：对于其中的一个点*i*来说：计算  $a(i) = \text{average}(i \text{ 向量到所有它属于的类别中其它点的距离})$ ，表示*i*向量到同一簇内其他点不相似程度的平均值。计算  $b(i) = \min(i \text{ 向量到与它相邻最近的一个类别内的所有点的平均距离})$ ，表示*i*向量到其他簇的平均不相似程度的最小值。

$$S(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

2.混淆矩阵：

混淆矩阵类别评价指标

True Positive: 真正类(TP), 样本的真实类别是正类, 模型预测成为正类

False Negative: 假负类(FN), 样本的真实类别是正类, 模型预测成为负类

False Positive: 假正类(FP), 样本的真实类别是负类, 模型预测成为正类

True Negative: 真负类(TN), 样本的真实类别是负类, 模型预测成为负类

		Predicted condition		
		positive	negative	
True condition	positive	True Positive	False Negative (预测错误)	Recall=TP/(TP+FN)
	negative	False Positive (预测错误)	True Negative	FPR=FP/(FP+TN)
		$F = \frac{2 * P * R}{P + R}$	$P = \frac{TP}{TP + FP}$	Accuracy= $\frac{(TP + TN)}{(TP + FN + FP + TN)}$
			$FOR = \frac{FN}{(FN + TN)}$	

Recall: 召回率, 模型预测为正类的样本的数量, 占总的正类样本数量的比值。一般情况下, Recall越高, 说明有更多的正类样本被模型预测正确, 模型的效果越好。

FPR: 模型预测为正类的样本中, 占模型负类样本数量的比值。一般情况下, 假正率越低, 说明模型的效果越好。

Accuracy: 模型的精度, 即模型预测正确的个数/样本的总个数。一般情况下, 模型的精度越高, 说明模型的效果越好。

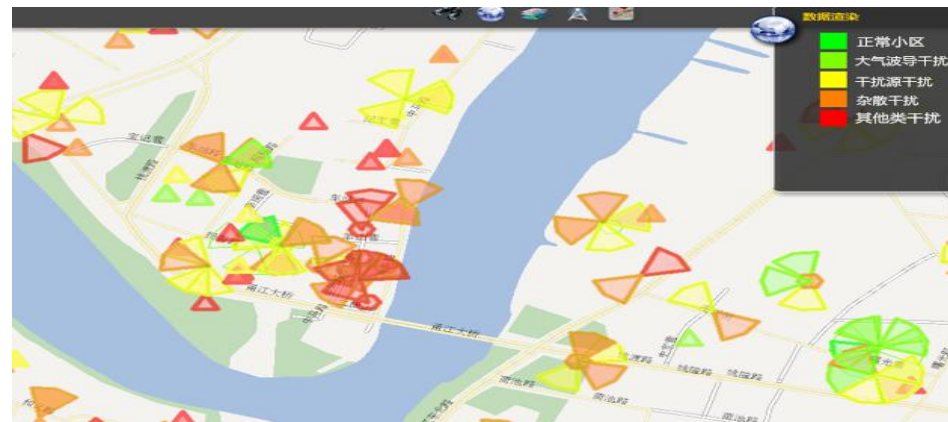
F: 综合了Precision和Recall的一个判断指标, F1-Score的值是从0到1, 1是最好, 0是最差

Precision: 查准率, 在模型预测为正类的样本中, 真正为正类的样本所占的比例。一般情况下, 查准率越高, 说明模型的效果越好。

FOR: 表示在模型预测为负类的样本中, 真正的正类所占的比例。一般情况下, 错误遗漏率越小, 模型的效果越好。

CLASS_DETAIL	CLASS	CGI	BANDWID	ULMEANN	ULMEANN	ULMEANN	ULMEANN	ULMEANN	ULMEANN	ULMEANN	ULMEANN	ULMEANN
0 其它干扰	460-00-1020560-131	10	-105.285	-105.297	-104.768	-104.764	-104.949	-105.174	-105.348	-105.455		
1 干扰器干扰	460-00-1022921-133	10	-107.48	-107.496	-107.178	-107.137	-107.045	-107.084	-107.082	-107.11		
2 互调谐波干扰	460-00-1023022-132	10	-112.068	-112.179	-109.055	-109.089	-109.387	-109.999	-110.567	-110.881		
3 业务同频弱干扰	460-00-1023022-133	10	-108.179	-108.213	-106.155	-105.949	-105.909	-106.114	-106.872	-107.175		
4 杂散干扰	460-00-1023022-134	10	-106.347	-106.5	-103.579	-102.566	-103.243	-105.027	-107.016	-108.77		
5 阻塞干扰	460-00-1023078-133	10	-104.8	-104.802	-104.931	-105.427	-105.611	-105.585	-105.967	-106.335		
6 大气波导干扰	460-00-1023081-133	10	-104.289	-104.476	-105.036	-105.738	-105.905	-105.57	-105.709	-105.676		
7 GPS干扰	460-00-1023085-132	10	-110.615	-110.468	-103.181	-89.3842	-100.51	-109.9	-112.123	-112.946		
8 广电MMDS干扰	460-00-1023191-134	10	-81.4751	-81.4512	-81.3834	-81.615	-82.1207	-82.7877	-83.3512	-83.5457		

干扰指纹库



页面呈现不同干扰





中国移动  
China Mobile

| 5G<sup>+</sup>创新实训基地  
技术实操的练兵场 · 能力认证的人才站 · 5G应用的孵化器

# 谢谢!



# 5G++ 创新实训基地

技术实操的练兵场 · 能力认证的人才站 · 5G 应用的孵化器

