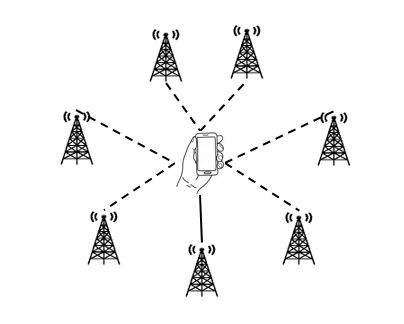
# 手机基站测量报告数据说明文档



## 数据说明

一台可以正常通信的手机，每时每刻都保持着和附近基站的连接，这些连接的信息组成了测量报告数据，也就是MR数据。MR数据中主要包括: 相连基站的ID以及对应的信号测量信息。它们和手机的上下文信息共同组成了MR数据。一条典型的MR数据格式如下：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| TrajID | IMEI | IMSI | MRTime |  |
| RNCID\_1 | CellID\_1 | Dbm\_1 | AsuLevel\_1 | SignalLevel\_1 |
| ... | ... | ... | ... | ... |
| RNCID\_6 | CellID\_6 | Dbm\_6 | AsuLevel\_6 | SignalLevel\_6 |
| Longitude | Latitude | Altitude | Accuracy | Speed |

其中，TrajID为轨迹ID， IMEI和IMSI共同组成手机的唯一ID，MRTime为当前时间戳；RNCID和CellID共同组成基站的全球唯一ID，对应的基站GPS坐标可以在工参表中查询得到，Dbm，AsuLevel和SignalLevel是手机和基站间的信号测量数据；Longitude和Latitude是GPS位置标签，手机在报告位置同时，也会给出定位的理论误差Accuracy以及理论高度Altitude；Speed是当前的运动速度，单位m/s。

注意，数据中存在缺失值，以-999和-1的形式存在。

GPS标签的位置范围是：左下角(31.28175691, 121.20120485)，右上角(31.29339344, 121.21831882)，可从数据中统计得到。若转化成UTM坐标，则：左下角(328770, 3462224)，右上角(330420, 3463487)。可根据计算需要适当扩大范围。

\*\*度量指标\*\* 即预测位置与真实位置之间的误差，单位是米，可以先将GPS坐标转化成UTM坐标再使用欧氏距离，也可以使用地球半径等参数根据GPS坐标计算。使用平均误差，中位误差，90%误差等度量。

\*\*地图投影\*\* 可以使用百度地图，谷歌地图，OpenStreetMap等提供的API投影。注意它们可能使用不同的坐标系。

注意模型最终loss并不能作为评估指标。

## 提交说明

train2g.csv中给出训练数据，test2g.csv测试数据不带标签，gongcan.csv给出基站GPS信息。

对于每一小问，输出pred.csv文件，每一行对应test2g.csv中每一条MR数据的预测GPS位置(逗号分隔，可参考样例输出)。

提交格式为GroupID\_TeamLeaderID.zip，目录结构：

GroupID\_TeamLeaderID:

a

src/

pred.csv

b

src/

pred.csv

...