

工参治理的主要原则

- 导入工参数据，校验基本逻辑错误
- 通过MDT数据对4G基站经纬度纠偏
- 通过关联业务量数据对应删未删、应加未加等工参问题做筛查

流程

数据入库

工参入库

- 数据库中经纬度、方向角、挂高、下倾角都是数值格式的，要将这些字段转成数值，尤其是4G工参，非数值格式的比较多，目前的操作是将下列字符替换成空：/* 和*和#N/A。还要将o（字母欧）替换成0（数字零），最后再检查看有没有其他非0值
- csv是逗号分隔的格式，有的小区名中含逗号，需要将逗号替换成空
- 字段顺序对应上，保存成utf-8格式
- 执行存储过程rpt_0_4g_siteinfo_backup和rpt_0_5g_siteinfo_backup，备份下工参数据
- 将上面的新的工参数据导入数据库cfg_0_4g_siteinfo和cfg_0_5g_siteinfo

业务量入库

- 业务量的数据源目前是华为直接推送，本地41服务器：/data/HUAWEI/SDR_FLOW_CELL/，
- 目前已自动入库，入库脚本位置，本地09：/home/do/jyc/sdrflow/sdrflow_ftp.py
- 数据量较大需要分区，分区程序是调用存储过程sqmdb_bak.huadan_partition()。脚本位置，本地09：/home/do/jyc/sdrflow/partition_sdr.sh
- 入库以后需要汇聚，汇聚程序位置，本地09：/home/do/jyc/sdrflow/gczl.sh。汇聚结果每天4G约33万行，5G约8万行。没有分区。

MDT数据处理

- 每次更新前找郝建输出近期的MDT数据到集团hive：fastdo_lte.hb_mdt_grid_cell_manual_zl
- 提取12各地市的MDT数据脚本如下：

```
citys=(BAODING CANGZHOU CHENGDE HANDAN HENGSHUI LANGFANG QINHUANGDAO
SHIJIAZHANG TANGSHAN XINGTAI XIONGAN ZHANGJIAKOU);
for city in ${citys[@]};
do
beeline -u "jdbc:hive2://10.168.26.1:10000/fastdo_lte;" -n zhihui030 -p
mAGR_H7b2u_Gbi7N --outputformat=csv2 --showHeader=true -e "select
sdate,city,enbid,eci,gridx,gridy,rsrp_samples from
fastdo_lte.hb_mdt_grid_cell_manual_zl where ct='$city' and rsrp_samples>5 and
ds='20220222'" > $city'.csv';
done
```

- 对MDT数据做聚类处理，程序位置，本地137：
C:\Users\zhanglei\PycharmProjects\untitled2\dbscan.py
- 将处理结果入库，源数据入库到o2p.sdr_0_4g_mdt_g50_cell_day，处理后的结果入库到o2p.sdr_0_4g_mdt_g50_dbscan_cay

数据处理和输出 (Oracle转pg后, 本流程还未验证, 预计3月21日验证完毕)

创建视图 (如果需求没有变动, 创建1次视图就可以)

执行下面的存储过程

perform_4g_sdrflow

perform_5g_sdrflow

rpt_0_4g_siteinfo_error

rpt_0_5g_siteinfo_error

rpt_0_4g_mdt_deviate_db

rpt_0_4g_flow_no_siteinfo

rpt_0_5g_flow_no_siteinfo

输出结果

执行接口文件程序输出接口文件

执行刘子龙需求程序输出对应文件

其他

- pg库连接方式:

地址: 192.168.0.9

端口: 5432

数据库: sqmmt

账号/密码: pmparse/mdasil

- 算法说明见文件《工参治理资源大屏总结 V1 20210817.pdf》 (张磊)