# Cm数据采集

脚本：cm\_collect.py

配置：

目录：

输入：

ftp文件 && telnet 返回

输出：

ftp：  
cm\_lte\_enbfunction\_hw\_bak  
cm\_lte\_eutrancelltdd\_hw\_bak  
cm\_lte\_eutranrelationtdd\_hw\_bak  
  
telnet：  
cm\_lte\_dsp\_cell\_hw\_bak  
param\_lte\_hw\_login\_bak  
cm\_lte\_lst\_cagroupcell\_hw\_bak  
cm\_lte\_lst\_cell\_hw\_bak  
  
rst\_l2l\_mr\_detail

# Pm数据采集

脚本：pm\_collect.py

配置：

目录：

输入：

ftp文件

输出：

pm\_lte\_eutrancelltdd\_hw

# Mro数据采集

脚本： mro\_collect\_lte\_hw.py

配置：

目录：

输入： ftp文件

输出：

pm\_lte\_mro\_hw

# Mro文件解析

解析已经下载的mro文件

1. 根据正则表达式匹配指定目录下的mro文件，并生成scv文件，每个文件500行，每一行都是一个mro文件名称
2. 使用LteHWCellCollectOutPutHandle解析文件并生成csv文件
3. 将csv文件数据导入到表中，LTE\_NCELL\_INFO 格式的文件数据导入到rst\_l2l\_hw\_hz\_sm表中，其他格式的文件数据导入到rst\_l2l\_hw\_nc\_sm表中

脚本：test\_mr\_paser\_sm.py

配置：

目录：

输入：

Mro ftp文件

输出:

LTE\_NCELL\_INFO 格式的文件🡪 rst\_l2l\_hw\_nc\_sm <目前使用的>

rst\_pm\_l2l\_hw\_hz  （脚本没有携带表后缀时）

非LTE\_NCELL\_INFO 格式的文件🡪 rst\_l2l\_hw\_hz\_sm<目前使用的>

rst\_pm\_l2l\_hw\_nc\_info （脚本没有携带表后缀时）

* **解析生成的****rst\_l2l\_hw\_hz\_sm将在重叠覆盖计算时用于生成rst\_l2l\_mro\_hw表数据。**

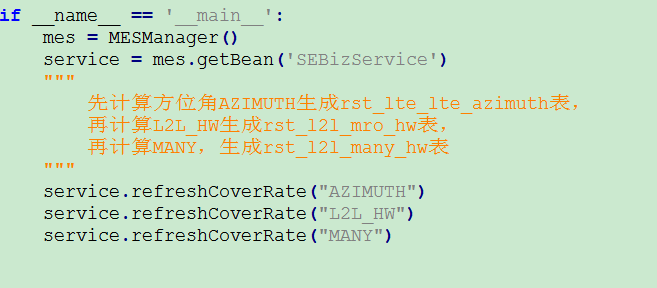
# 重叠覆盖度计算

概述

重叠覆盖度计算需要先计算方位角，在计算L2L\_HW生成rst\_l2l\_mro\_hw，最后计算MANY，生成rst\_l2l\_many\_hw表，因为计算MANY的时候需要依赖rst\_l2l\_mro\_hw数据。

基于mr的重叠覆盖计算根据邻区关系信息和mr数据计算筛选g2g,t2g,l2t,l2l覆盖情况；

基于方位角的重叠覆盖计算，根据小区的工参数据，通过经纬度计算两个小区的距离，筛选出小区距离小于50米的共站小区，输出g2g共站列表，t2g共站列表，l2t共站列表和l2l共站列表，再从四个共站列表中筛选，方位角偏差10度的为补偿小区。



流程说明：

需要先计算一补一再计算多补一的情况，根据数据的前后依赖关系配置脚本的执行顺序。

1. 通过groupname 查询mes\_over\_setting中的表数据，groupname为空时，先查询一补一再查询多补一
2. 遍历上一步的查询结果
3. 通过mes\_over\_setting.querysql查询数据并生成临时表，如果数据超过80000条则分批处理
4. 使用handle处理（简单计算）querysql查询到的数据并将数据写入到csv中
5. 将写入到csv文件的数据导入到mes\_over\_setting.res\_table中

脚本：mro\_over\_cal.py

配置：mes\_over\_setting

目录：

输入：

mes\_base\_info\_lte  🡪用于计算方位角生成rst\_lte\_lte\_azimuth数据

**rst\_l2l\_hw\_hz\_sm**  🡪 用于mr生成 rst\_l2l\_mro\_hw

rst\_l2l\_mro\_hw，**rst\_l2l\_hw\_hz\_sm 🡪** **rst\_l2l\_many\_hw**

输出：

rst\_lte\_lte\_azimuth

rst\_l2l\_mro\_hw

rst\_l2l\_many\_hw

# 大数据预测功能

根据不同制式的pm\_gsm|td|lte\_\*\_his历史数据生成 rst\_bigdata\_pm开头的预测表数据。核心思路就是通过过去30天同一时段的pm均值来预测明天同一时段的pm值。

1. 查询mes\_bigdata\_forecast\_setting表配置，按需要配置
2. 组装过去同一时刻时间值，如’2018-08-29 00:15:00’, ’2018-08-30 00:15:00’……
3. 将时间作为条件查询pm\_lte\_eutrancelltdd\_hw\_his表数据获取均值
4. 生成文件，并将文件导入到mes\_bigdata\_forecast\_setting.res\_table 中

脚本：bigdata\_forecast.py

配置：

mes\_bigdata\_forecast\_setting.lte

目录：

输入：

pm\_lte\_eutrancelltdd\_hw\_his

输出：

rst\_bigdata\_pm\_lte\_eutrancelltdd\_hw

* **大数据预测功能生成的rst\_bigdata\_pm\_lte\_eutrancelltdd\_hw表数据将在计算历史Kpi时被用到**

# 历史Kpi计算

通过rst\_bigdata\_pm\_lte\_eutrancelltdd\_hw , cm\_lte\_eutrancelltdd\_hw, cm\_lte\_enbfunction\_hw 表数据的关联生成mes\_lte\_kpi表数据，该脚本在代码里面已经写死，只能在0~6点之间执行，其他时间都不会执行。

脚本：bigdata\_next.py

配置：mes\_kpical\_set.BIGDATA\_KPI\_CAL\_HIS

目录：

输入：

rst\_bigdata\_pm\_lte\_eutrancelltdd\_hw JOIN  
cm\_lte\_eutrancelltdd\_hw JOIN  
cm\_lte\_enbfunction\_hw

输出：

mes\_lte\_kpi

* **该步骤输出的mes\_lte\_kpi表数据将在休眠小区筛选时用到。**

# Real Kpi计算

脚本：pm\_collect.py

配置：mes\_kpical\_set.KPI\_CAL\_REAL

目录：

输入：

pm\_lte\_eutrancelltdd\_hw JOIN

cm\_lte\_eutrancelltdd\_hw JOIN

cm\_lte\_enbfunction\_hw

输出：

mes\_lte\_real\_kpi

# 休眠小区筛选

筛选休眠小区：

1. 查询busytype值，该值由mes\_zs\_coverage.base\_overlay\_degree确认，如果base\_overlay\_degree=1则为true，否则为false
2. 如果busytype值为true，则group\_name = azimuth\_l2l,为false则group\_name = mr\_l2l
3. 使用上一步骤得到的group\_name值作为参数查询mes\_sleepsel\_setting表数据
4. 如果mes\_sleepsel\_setting.del\_flag = 1 则先删除mes\_sleepsel\_setting.res\_table数据
5. 通过mes\_sleepsel\_setting.querysql查询数据
6. 通过mes\_sleepsel\_setting.service\_handle处理数据
7. 生成csv文件
8. 将csv文件数据导入到mes\_sleepsel\_setting.res\_table中

脚本： pm\_collect.py

配置： mes\_sleepsel\_setting.'%l2l'

目录：

输入：

rst\_l2l\_energyref\_azimuth JOIN  
mes\_lte\_kpi  
  
rst\_l2l\_energyref\_mr JOIN  
mes\_lte\_kpi  
  
rst\_l2l\_energyref\_many JOIN  
mes\_lte\_kpi

输出：

mes\_l2l\_sleep\_azimuth  
  
mes\_l2l\_sleep\_mr  
  
mes\_l2l\_sleep\_many

mes\_l2l\_mobile\_table;

# 休眠小区冲突处理

由于小区既有可能是休眠小区，又可能是补偿小区业务场景的存在，导致，需要对小区休眠做一个优先级，尤其是在不同制式的场景下，如小区GSM-A是 TD-B的休眠小区，又是Lte的补偿小区时，应该把GSM优先作为休眠小区而不是做为补偿小区，通常情况下休眠优先级从高到低依次是GSM>TD>LTE。

相同制式，小区是休眠小区，也是补偿小区，则该小区只做休眠小区，不做补偿小区考虑，需要将作为补偿小区的一组数据删除，避免被补偿的小区休眠导致业务异常。

# 休眠指令生成

脚本：pm\_collect.py

配置：mes\_sleepexe\_setting.'l2l'

目录：

输入：

mes\_l2l\_sleep\_azimuth  
mes\_l2l\_sleep\_many  
mes\_l2l\_sleep\_mr  
  
cm\_lte\_powersaving\_zte  
  
cm\_lte\_dsp\_cell\_hw  
param\_lte\_hw\_login  
cm\_lte\_lst\_cagroupcell\_hw  
mes\_l2l\_currentsleep

输出：

mes\_adjust\_command

# 休眠指令下发/唤醒

脚本：

配置：

目录：

输入：

mes\_adjust\_command

输出：

休眠成功 mes\_lte\_commandlog  
  
休眠失败 mes\_td\_black  
 mes\_subscribe\_lte\_selfalarm

# 历史数据预测 8:00

每天8:00计算能时段（ 0:00-6:00）的业务模型，根据 cm\_\* 表和pm\_\*\_his表生成rst\_\* 表。

根据\*表的数据量进行自动适配查询三个月，一个月还是七天的数据。



根据历史数据的存储量使用不同的计算公式，如果系统有的n月（周）的数据则:

满足三个月历史数据(n > 3m)：

下一时段预测业务量=前第一个月同期下一时段业务流量均值\*0.3+前第二个月同期下一时段业务量均值\*0.2+前第3至6个月同期下一时段业务量均值\*0.1

满足2个月不到三个月(2m < n < 3m):

下一时段预测业务量=前第一个月同期下一时段业务流量均值\*0.3+前第二个月同期下一时段业务量均值\*0.2

满足1个月不到2个月(1m < n < 2m):

下一时段预测业务量=前第一个月同期下一时段业务流量均值\*0.4

满足一周不满足一个月(1w < n < 1m):

下一时段预测业务量=前第一个月同期下一时段业务流量均值\*0.4

不满足一周 (n < 1w):

下一时段预测业务量=前第一个月同期下一时段业务流量均值\*0.5

脚本：his\_fcast.py

配置：mes\_fcast\_setting

目录：

输入：

cm\_gsm\_cellgprs\_eric && pm\_gsm\_cellgprs\_eric\_his

cm\_gsm\_cellqoseg\_eric && pm\_gsm\_cellqoseg\_eric\_his

cm\_gsm\_celtchf\_eric && pm\_gsm\_celtchf\_eric\_his

cm\_gsm\_cellqosg\_eric && pm\_gsm\_cellqosg\_eric\_his

cm\_gsm\_celtchfp\_eric && pm\_gsm\_celtchfp\_eric\_his

cm\_gsm\_celtchh\_eric && pm\_gsm\_celtchh\_eric\_his

cm\_gsm\_cltch\_eric && pm\_gsm\_cltch\_eric\_his

cm\_gsm\_mrf\_4c0003ed\_hw && pm\_gsm\_mrf\_4c0003ed\_hw\_his

cm\_td\_carrier\_hw\_cell && pm\_td\_carrier\_hw\_cell\_his

cm\_td\_carrier\_zte\_cell && pm\_td\_carrier\_zte\_cell\_his

cm\_td\_utrancell\_hw && pm\_td\_utrancell\_hw\_his

cm\_td\_utrancell\_zte && pm\_td\_utrancell\_zte\_his

cm\_lte\_eutrancelltdd\_hw && pm\_lte\_eutrancelltdd\_hw\_his

cm\_lte\_eutrancelltdd\_zte && pm\_lte\_eutrancelltdd\_zte\_his

cm\_lte\_eutrancelltdd\_eric && pm\_lte\_eutrancelltdd\_eric\_his

cm\_gsm\_trafdlgprs\_eric && pm\_gsm\_trafdlgprs\_eric\_his

输出：

rst\_pm\_gsm\_cellgprs\_eric

rst\_pm\_gsm\_cellqoseg\_eric

rst\_pm\_gsm\_celtchf\_eric

rst\_pm\_gsm\_cellqosg\_eric

rst\_pm\_gsm\_celtchfp\_eric

rst\_pm\_gsm\_celtchh\_eric

rst\_pm\_gsm\_cltch\_eric

rst\_pm\_gsm\_mrf\_4c0003ed\_hw

rst\_pm\_td\_carrier\_hw\_cell

rst\_pm\_td\_carrier\_zte\_cell

rst\_pm\_td\_utrancell\_hw

rst\_pm\_td\_utrancell\_zte

rst\_pm\_lte\_eutrancelltdd\_hw

rst\_pm\_lte\_eutrancelltdd\_zte

rst\_pm\_lte\_eutrancelltdd\_eric

rst\_pm\_gsm\_trafdlgprs\_eric

# Busytype 值说明

mes\_zs\_coverage 表中base\_overlay\_degree = 1时Busytype=true，否则 Busytype=false

对于只有4g的场景下，建议只打开l2l的场景。

# rst\_l2l\_energyref\_\*生产规则

SELECT \* FROM mes\_celllist\_ref a WHERE a.`res\_table` LIKE 'rst\_l2l\_energyref\_%' AND a.`enabled` = 1;

查询出来的sql就是rst\_l2l\_energyref\_ 表对应的数据生成sql

# rst\_bigdata\_pm\_\* 表数据生成规则

所有的rst\_bigdata\_pm打头的表都是有bigdata\_forecast.py运行生成的。

# rst\_\*\_mr\_detail 表生成规则

通过rst\_\*\_mr\_detail 格式的表都是通过mes\_mrdetail\_model配置生成的，都是在执行cm\_collect.py脚本时通过service.improveMrData("")生成的。如果只有l2l场景则只需要执行service.improveMrData("l2l")即可。

# rst\_lte\_lte\_azimuth

rst\_lte\_lte\_azimuth 表数据由mes\_base\_info\_lte 自己关联后得到数据，数据生成规则配置为

SELECT \* FROM mes\_over\_setting a WHERE a.`result\_table` = 'rst\_lte\_lte\_azimuth';

rst\_l2l\_mro\_hw 数据由 rst\_l2l\_hw\_hz\_sm 生成

rst\_l2l\_many\_hw 数据由rst\_l2l\_hw\_hz\_sm a LEFT JOIN rst\_l2l\_mro\_hw 生成。

# mes\_fcastnext\_setting 作用？

SELECT \* FROM mes\_fcastnext\_setting a WHERE a.`source\_table` LIKE '%hw%' AND a.`source\_table` LIKE '%lte%';

# mes\_fcast\_setting 作用说明

用于历史数据预测，生成rst\_ 打头的表。

# 脚本作用说明

alarm.py

alystest.py

azimuth.py

azimuth\_converge.py

bigdata\_forecast.py

根据过去30天的pm数据预测下一天的pm数据生成rst\_bigdata\_pm 打头的表

bigdata\_next.py

生产历史Kpi数据,mes\_lte\_kpi数据

cell\_sleep\_time.py

checkcmdserver.py

cleanHisData.py

清楚历史数据

cm\_collect.py

cm 数据采集解析&& rst\_\*\_mr\_detail数据生成

his\_fcast.py

生成rst\_pm\_ 相关的表数据

mro\_collect\_\*.py

mro数据采集

mro\_over\_cal.py

计算重叠覆盖度，生成方位角和mr覆盖数据

notifyall.py

唤醒所有数据

parameters.py

提供几个公用方法

performance\_cal\_day.py

按天，小时，分钟维度计算kpi数据

pm\_collect.py

pm采集解析，休眠小区的筛选，监控，休眠，唤醒

test\_mr\_paser\_sm.py

mro数据解析，test\_mr\_paser\_\*.py的脚本都是用于解析mro，但是其他的几类暂时没有发现有什么用

# 还无法解释的坑

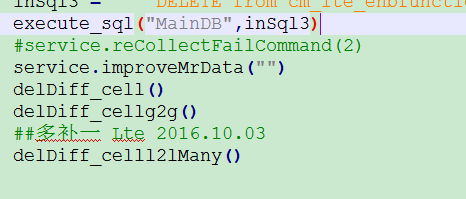
这里记录的都是目前还没有搞清楚的问题。

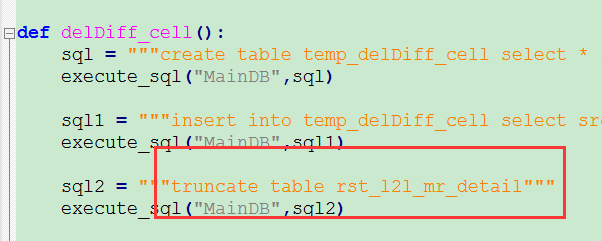
## 数据计算规则

在代码里面涉及到数据计算规则不清楚，可参考的资料是之前的概要设计里面提到的几个公式。

## rst\_\*\_mr\_detail 数据先写入再删除

在cm\_collect.py脚本中，先执行service.improveMrData("") 通过mes\_mrdetail\_model配置将数据写入到了rst\_\*\_mr\_detail 表中，随后又执行了delDiff\*()方法将数据全部删除。





# Mr，CM，PM数据作用及采集频率

基于历史的mr数据筛选可休眠的小区，筛选出来的小区列表就是系统具备休眠条件的小区，一般来说这些都是相对固定的，即系统可休眠的小区基本上是固定的。是否休眠需要根据pm的实时数据来决定，如果小区当前pm指标出现异常则不能休眠小区。Cm为小区的工参数据，下斜角，载频等会经常发生变化，但是基本信息不变。

但是由于MR数据特别大，需要每天采集前一天的数据，月末再执行重叠覆盖度计算逻辑。由于MR数据的生成多数都存在很大的延时问题，故我们将前一天MR数据的采集时间定为第二天13点采集前一天的数据，即当天13点采集前一天全天的mr数据。

PM（话统）数据只需要采集节能时段的pm数据及可。

Cm（工参）：为保证小区基础工参的准确性，通过mml（telnet）指令和ftp文件方式采集，每天22点提前采集小区的工参数据，第一次部署为实时采集。

告警数据：用于逻辑判断，每15分钟检查一次最新的告警数据。

MR数据采集为每天采集前以前的，计算是按月为一周期；历史业务模型预测周期为每天，时间为每天8点对历史进行预测，生产历史kpi数据；每天23点更新小区节能状态表，结合黑白名单对小区状态进行监控；



主逻辑流程分三部分：月粒度执行部分、天粒度执行部分、15分钟粒度执行部分，具体流程如下：

按月周期：

1、按月为周期，更新GSM/TD-SCDMA/LTE各网制式及设备的MR数据，如果没有MR数据的根据手动跟新粒度更新工参数据。

2、判断GSM/TD-SCDMA/LTE各网络制式及设备是否有MR采集数据，如果是有MR数据，则基于采集的MR数据结合邻区关系进行各网络对应的重叠覆盖度的计算；如果否没有MR数据，则基于手动提供的小区工参基础信息进行各网络的重叠覆盖度计算；

3、根据前面的重叠覆盖度计算结果，筛选输出GSM/TD-SCDMA/LTE各网络制式的节能小区表，节能小区包括休眠小区和补偿小区及对应关系；

按天周期：

4、按天为周期，在每天08:00开始对历史业务统计模型进行预测，根据存储的历史数据对节能时段（00:00-06:00）每个15分钟时段业务进行历史预测；

5、在每天23:00更新节能小区状态表，确保GSM/TD-SCDMA/LTE各网络小区状态表和现网运行和黑白名单状态保持一致；

15分钟周期：

6、15分钟粒度为周期，（1）每个周期开始时读取小区状态表和黑白名单；（2）同时对GSM/TD-SCDMA/LTE各网小区15分钟粒度监控小区KPI和告警，并根据监控情况实时处理告警、差小区、以及补偿唤醒小区流程；

分支一：

7、根据读取GSM/TD-SCDMA/LTE各网络小区状态和黑白名单信息，对于上次周期结果列表中开启的小区进行各网络休眠和补偿小区筛选；输出各网络的休眠小区和补偿小区列表；

8、根据各网络制式的节能优先级设置情况，对各网络的休眠小区和补偿小区冲突处理，确定优先节能的制式小区，消除既是休眠小区有作为补偿小区情况，最后输出可执行的休眠小区和对应的补偿小区列表。

9、根据最后输出的可执行的休眠小区和补偿小区信息列表，对休眠小区进行休眠操作执行。

分支二：

10、根据读取GSM/TD-SCDMA/LTE各网络小区状态和黑白名单信息，对于上次周期结果列表中关闭的小区实时监控其补偿小区；

11、如果监控到补偿小区有告警或业务负荷超过设置门限，则根据相应的唤醒规则对选择对应的唤醒小区，输出唤醒小区列表信息；

12、对唤醒小区列表信息执行唤醒操作；

合并：

13、根据本15分钟时段执行休眠和唤醒操作以及全网小区监控结果，更新节能小区状态表和黑白名单（黑名单添加需满足命令执行失败小区和连续N次出现指标差和高级别告警的小区，N默认为3次）。

# Pm数据prb利用率为0 ？

# Pm，cm，mr数据处理流 ？

那些任务使用到了pm，cm，mr数据，每个任务数据处理的目标和作用。

# cm\_lte\_dsp\_cell\_hw 数据来源及作用？

cm\_lte\_dsp\_cell\_hw表里面存储的数据是在哪个步骤里面出现的，数据的作用，维护和使用该数据的环节。

# 数据验证 ？

Mr重叠覆盖数据，方位角覆盖数据，休眠小区筛选数据 验证问题