# **场客流量的时空分布预测**

# Oh my god团队

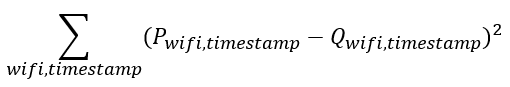
## 赛题背景

机场拥有巨大的旅客吞吐量，人员流动呈现**高动态**，**高密度可变**，**时间空间分布不均匀**等特点，与之相对应的则是巨大的服务压力。安防、安检、突发事件应急、值机、行李追踪等机场服务都希望能够预测未来的旅客吞吐量，并据此提前调配人力物力，更好的为旅客服务。本次大赛提供了广州白云机场9月10号到11月10号近两个月的真实数据记录（WIFI连接记录，安检旅客记录，航班排班表等），要求参赛团队构建客流量预测模型预测未来两个整天机场内每个WIFI\_AP点每十分钟的内的平均设备连接数量。

## 赛解解读

2.1目标解读

设P为选手提交旅客流量预测值，Q为机场WIFI\_AP实测的旅客流量。



通过对目标函数的理解，主要是缩小预测值和真实值的均方误差。那么重点自然是客流量较大且不稳定的区域。

2.2数据分析

首先通过对对不同该区域进行可视化比较，分析得到不同区域的客流量的均值和波动性有较大的差别。初步对比分析之后可以得到如下结论：

* 1. 登机区域的客流量平均值，与走廊区域和航站楼区域相比要高出很多，且存在比较大的波动性；
  2. 通过统计WIFI客流量均值水平，可以看出均值小于2的WIFI数占到了总数40%左右，且较小的WIFI点多分布在走廊区域和航站楼区域；
  3. 通过历史统计值来拟合的客流量有一定的参考意义，尤其是针对客流量平均值较小的WIFI点。
  4. 登机区域客流量波动的产生与航班的起飞时间关联性较大；

综上，我们将分开建模，分别为基于历史统计信息的整体区域模型以及基于航班分析的登机区域模型；分开建模的主体思想在于考虑到登机区域直接受到航班的影响，通过建立航班和登机区域的关联，能够比较准确的预测登机口区域的客流量波动情况。而对于整体区域来说，客流量分布随日期变化时段变化较均匀，与航班信息的联系略弱，同时考虑分开建模，一定程度上增大了模型间的多样性，经过模型间的融合可以更好的提升预测准确率；

2.3 竞赛过程

1. 分析赛题背景及预测目标，了解实际业务特点；
2. 通过可视化及统计分析及数据探索得到业务逻辑，数据分布特点；
3. 建立整体区域规则，建立统计平均水平预测，并指导特征工程建立；
4. 结合业务分析和整体区域规则建立整体区域回归预测模型；
5. 通过特征选择，训练日期划分，选取不同的训练基模型并融合规则算法得到整体区域预测模型；
6. 分析登机口区域及WIFI点的联系情况；
7. 分析登机口区域WIFI点的业务波动情况与航班信息的关联情况；
8. 分析整理客流人群的安检习惯，整合航班影响范围内不同的时间段的影响因子；
9. 建立登机口区域的预测模型；
10. 融合整体区域模型和登机口区域模型得到最终模型架构；

## 核心思路

3.1 整体区域建模

3.1.1 影响因素分析

通过整体区域建模希望能够较为准确预测客流量的平均分布，并且比较好的预测一些不是由如航班信息等外部信息影响而造成的波动情况。通过对数据分析，得到如下能够影响客流量的因素：

历史wifi连接统计信息;

时段的影响;

星期对接入量的影响;

工作日，节假日的影响;

楼层，区域的影响

本方案中我们建立两种方法预测；一是规则方法用于把握业务的统计平均状态，并重点关注我们通过数据分析得到的主观业务测重点；二是采用机器学习方法，通过滑窗回归，更好的预测潜在的业务波动。

3.1.2规则方法的业务逻辑

首先思考影响客流量的业务逻辑，业务为本，人工规则的设计从希望能更好的反映实际业务特点出发，并一定程度指导特征工程，通过业务思考，得到如下有效的业务逻辑，并通过建立线下测试集合，找到最优加权参数；

* 去除存在数据缺失以及过节的日期数据；
* 明显高于或低于的均值水平的数据点应当予以平滑；
* 近期历史业务的统计均值具有较强的参考价值；
* 对于预测时段中接近已有数据的时段，历史数据指导意义更强；
* wifi接入量与星期成一定的相关性；
* 周五业务量较周六业务量稍大，且根据趋势判断线上预测业务量欠饱和；
* 1,2,3三层楼，W1,W2,W3,E1,E2,E3,EC,WC,T1九个区域，其业务量平均水平，随日期变化时段变化的波动情况在区域内存在相似性，区域楼层间存在一定的差异性。

3.1.3模型方法之特征工程

根据规则的业务逻辑，构建相关特征如下：

* wifi\_id、小时分钟、所处楼层区域、时段进行one-hot编码；
* 前1,3,5,7,14天对应十分钟，对应小时统计信息；
* 对应星期的统计特征；
* 对应区域的统计特征；
* wifi接入量及方差排名特征；
* 交叉特征等；

在构建完特征工程之后，采用GBDT，LR做为不同的基准模型，并通过采用不同特征子集划分，不同的训练集日期划分得到具有差异性的不同基模型预测结果，用于后续的模型融合；

3.2 登机区域建模

对于登机区域来说，客流量均值较大且波动情况较为复杂。通过业务背景分析，考虑产生波动性的主要原因。总结如下：

1. 构建登机口与WIFI点的对应关系

为方便提取WIFI点的特征，需要构建登机口与WIFI点的对应关系，具体如下：

* 登机区域分别为六个区域的三楼和一楼；
* 只考虑登机口左右两边的WiFi（由于水平步道，不考虑对面的wifi），为每个登机口对应最近的一个或两个WIFI；
* 根据所提供信息，对应了三楼和一楼的登机口，共计190个；

1. 航班计划情况

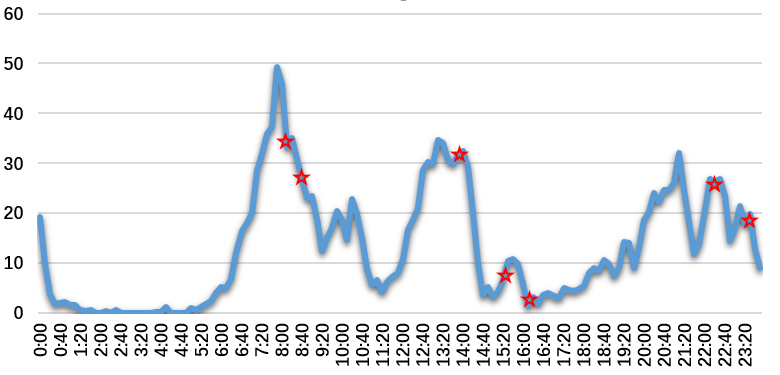


图1 连接量随时间变化图

航班计划情况是影响旅客动向的主要因素，一般在航班起飞之前，登机口比较容易汇聚较多的旅客；

1. 旅客安检数目的时刻分布

通过航班起飞能够估计波动出现的时刻，但是与波动幅度最为相关的应该是航班的载客量以及旅客提前多久进入登机区域。通过对旅客安检数据分析，得到旅客提前安检的时间统计，如图2所示。

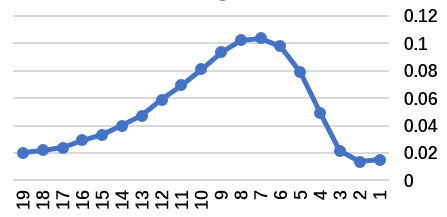


图2 安检提前时间统计

在得到旅客提前安检的分布，就可以拟合旅客到达登机口的汇聚情况。

1. 航班的载客量

通过统计安检历史数据，得到每次航班的平均值作为预测航班的载客量；

构建登机区域的特征工程

在得到上述四个数据之后，构建特征工程如下：

* Wifi点附近的登机口在未来不同时间阶段内的航班数目（时间段的选取为当前时间的之后半小时，以及之后半小时，一小时，...，三小时）；
* 对应航班历史乘客数目特征；
* 不同时间段所对应的影响因子；
* 其他特征，如Wifi点的历史统计量，时间特征等。

通过GBDT模型训练预测得到结果，与整体模型的结果融合；

3.3 整体框架

解决方案的整体框架如图3所示：

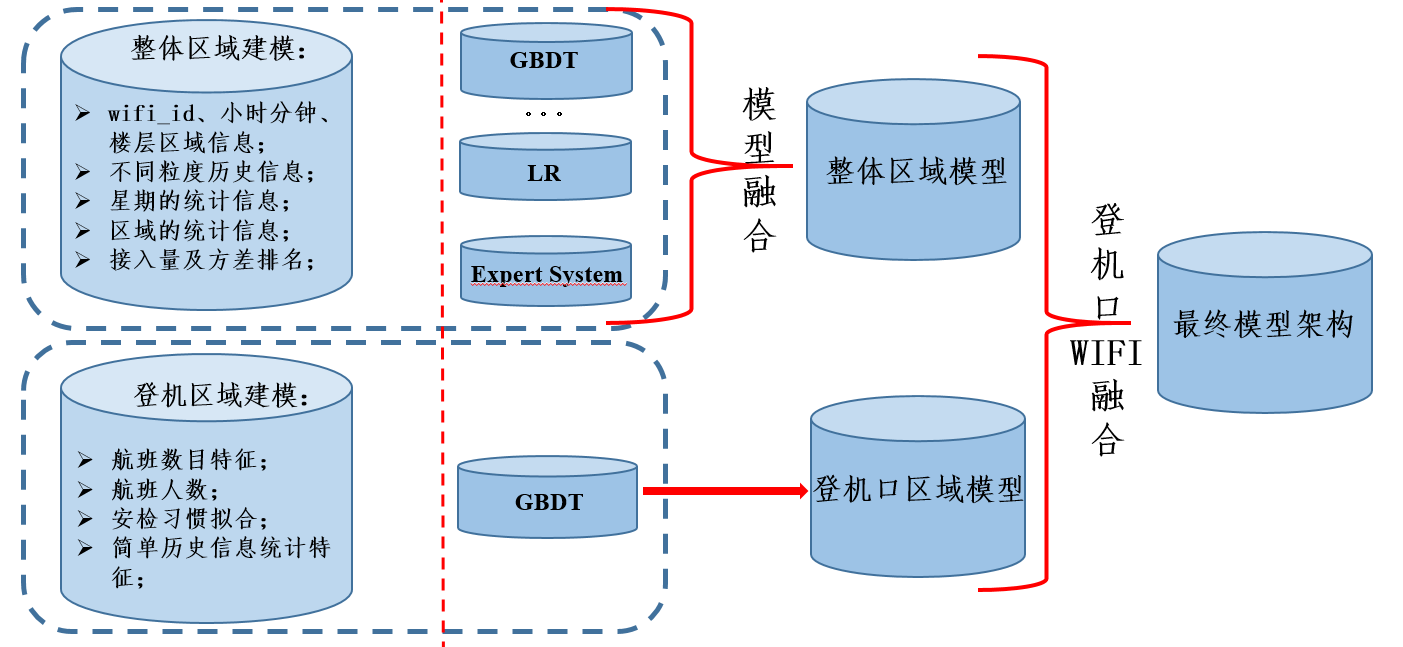


图3 解决方案整体框架

通过规则方法和滑窗回归方法得到稳定的整体区域的预测结果，并根据航班信息构建登机区域模型，能够准确的预测波动出现的时刻以及幅度。

## 关键代码

详见附录

## 赛后总结

* 结合目标对数据进行分析十分重要，方便赛题主要矛盾，突破点；
* 好的方法源于对数据的分析以及对业务的理解，而构造特征的过程就是刻画业务的过程；
* 团队合作让我们彼此学习的同时，也让我们体会到责任与担当；
* 比赛虽止但奋斗不止；

附录

4.1整体模型

特征提取，以线上测试集为例：

-------------------------------------------------------------------------

--对应分钟，近1,3,5,7,14天按照分段的均值，中位数，标准差；

drop table if exists AJ\_FeatureUsedWithSect\_11\_11;

create table if not exists AJ\_FeatureUsedWithSect\_11\_11

as select

t1.wifi\_ap\_tag,

t1.sect,

t1.hourTimeUsed,

t1.preDayBefMean,

t1.preDayBefMedian,

t1.preDayBefMean+t1.preDayBefMedian as preDayBefSumMeanMedian,

abs(t1.preDayBefMean-t1.preDayBefMedian) as preDayBefDiffMeanMedian,

t2.preThreeMean,

t2.preThreeMedian,

t2.preThreeStd,

t2.preThreeMean+t2.preThreeMedian as preThreeSumMeanMedian,

abs(t2.preThreeMean-t2.preThreeMedian) as preThreeDiffMeanMedian,

t3.preFiveMean,

t3.preFiveMedian,

t3.preFiveStd,

t3.preFiveMean+t3.preFiveMedian as preFiveSumMeanMedian,

abs(t3.preFiveMean-t3.preFiveMedian) as preFiveDiffMeanMedian,

t3.preFiveMin,

t3.preFiveMax,

t3.preFiveMax-t3.preFiveMin as preFiveDiffMaxMin,

t4.preSevenMean,

t4.preSevenMedian,

t4.preSevenStd,

t4.preSevenMean+t4.preSevenMedian as preSevenSumMeanMedian,

abs(t4.preSevenMean-t4.preSevenMedian) as preSevenDiffMeanMedian,

t4.preSevenMin,

t4.preSevenMax,

t4.preSevenMax-t4.preSevenMin as preSevenDiffMaxMin,

t5.preFourteenMean,

t5.preFourteenMedian,

t5.preFourteenStd,

t5.preFourteenMean+t5.preFourteenMedian as preFourteenSumMeanMedian,

abs(t5.preFourteenMean-t5.preFourteenMedian) as preFourteenDiffMeanMedian,

t5.preFourteenMin,

t5.preFourteenMax,

t5.preFourteenMax-t5.preFourteenMin as preFourteenDiffMaxMin

from

(

select

wifi\_ap\_tag,

sect,

hourTimeUsed,

avg(passenger\_count) as preDayBefMean,

median(passenger\_count) as preDayBefMedian

from

(

select wifi\_ap\_tag,sect,hourTimeUsed,passenger\_count

from AJ\_wifi\_ap\_slice10min

where dateTimeUsed in ("11-10")

)tb1

group by wifi\_ap\_tag, sect, hourTimeUsed

)t1

left outer join

(

select

wifi\_ap\_tag,

sect,

hourTimeUsed,

avg(passenger\_count) as preThreeMean,

median(passenger\_count) as preThreeMedian,

stddev(passenger\_count) as preThreeStd

from

(

select wifi\_ap\_tag,sect,hourTimeUsed,passenger\_count

from AJ\_wifi\_ap\_slice10min

where dateTimeUsed in ("11-08","11-09","11-10")

)tb2

group by wifi\_ap\_tag, sect, hourTimeUsed

)t2

on t1.wifi\_ap\_tag = t2.wifi\_ap\_tag and t1.sect = t2.sect and t1.hourTimeUsed = t2.hourTimeUsed

left outer join

(

select

wifi\_ap\_tag,

sect,

hourTimeUsed,

avg(passenger\_count) as preFiveMean,

median(passenger\_count) as preFiveMedian,

stddev(passenger\_count) as preFiveStd,

min(passenger\_count) as preFiveMin,

max(passenger\_count) as preFiveMax

from

(

select wifi\_ap\_tag,sect,hourTimeUsed,passenger\_count

from AJ\_wifi\_ap\_slice10min

where dateTimeUsed >= "11-06" and dateTimeUsed <="11-10"

)tb3

group by wifi\_ap\_tag, sect, hourTimeUsed

)t3

on t1.wifi\_ap\_tag = t3.wifi\_ap\_tag and t1.sect = t3.sect and t1.hourTimeUsed = t3.hourTimeUsed

left outer join

(

select

wifi\_ap\_tag,

sect,

hourTimeUsed,

avg(passenger\_count) as preSevenMean,

median(passenger\_count) as preSevenMedian,

stddev(passenger\_count) as preSevenStd,

min(passenger\_count) as preSevenMin,

max(passenger\_count) as preSevenMax

from

(

select wifi\_ap\_tag,sect,hourTimeUsed,passenger\_count

from AJ\_wifi\_ap\_slice10min

where dateTimeUsed >= "11-04" and dateTimeUsed <="11-10"

)tb4

group by wifi\_ap\_tag, sect, hourTimeUsed

)t4

on t1.wifi\_ap\_tag = t4.wifi\_ap\_tag and t1.sect = t4.sect and t1.hourTimeUsed = t4.hourTimeUsed

left outer join

(

select

wifi\_ap\_tag,

sect,

hourTimeUsed,

avg(passenger\_count) as preFourteenMean,

median(passenger\_count) as preFourteenMedian,

stddev(passenger\_count) as preFourteenStd,

min(passenger\_count) as preFourteenMin,

max(passenger\_count) as preFourteenMax

from

(

select wifi\_ap\_tag,sect,hourTimeUsed,passenger\_count

from AJ\_wifi\_ap\_slice10min

where dateTimeUsed >= "10-28" and dateTimeUsed <="11-10"

)tb5

group by wifi\_ap\_tag, sect, hourTimeUsed

)t5

on t1.wifi\_ap\_tag = t5.wifi\_ap\_tag and t1.sect = t5.sect and t1.hourTimeUsed = t5.hourTimeUsed;

----------------------------------------------------------

--近1,3,5,7,14天按照小时的均值，中位数，标准差；

drop table if exists AJ\_FeatureUsedWithHour\_11\_11;

create table if not exists AJ\_FeatureUsedWithHour\_11\_11

as select

t1.wifi\_ap\_tag,

t1.hourTimeUsed,

t1.preDayBefMeanHour,

t1.preDayBefMedianHour,

t1.preDayBefMeanHour+t1.preDayBefMedianHour as preDayBefSumMeanMedianHour,

abs(t1.preDayBefMeanHour-t1.preDayBefMedianHour) as preDayBefDiffMeanMedianHour,

t2.preThreeMeanHour,

t2.preThreeMedianHour,

t2.preThreeStdHour,

t2.preThreeMeanHour+t2.preThreeMedianHour as preThreeSumMeanMedianHour,

abs(t2.preThreeMeanHour-t2.preThreeMedianHour) as preThreeDiffMeanMedianHour,

t3.preFiveMeanHour,

t3.preFiveMedianHour,

t3.preFiveStdHour,

t3.preFiveMeanHour+t3.preFiveMedianHour as preFiveSumMeanMedianHour,

abs(t3.preFiveMeanHour-t3.preFiveMedianHour) as preFiveDiffMeanMedianHour,

t3.preFiveMinHour,

t3.preFiveMaxHour,

t3.preFiveMaxHour-t3.preFiveMinHour as preFiveDiffMaxMinHour,

t4.preSevenMeanHour,

t4.preSevenMedianHour,

t4.preSevenStdHour,

t4.preSevenMeanHour+t4.preSevenMedianHour as preSevenSumMeanMedianHour,

abs(t4.preSevenMeanHour-t4.preSevenMedianHour) as preSevenDiffMeanMedianHour,

t4.preSevenMinHour,

t4.preSevenMaxHour,

t4.preSevenMaxHour-t4.preSevenMinHour as preSevenDiffMaxMinHour,

t5.preFourteenMeanHour,

t5.preFourteenMedianHour,

t5.preFourteenStdHour,

t5.preFourteenMeanHour+t5.preFourteenMedianHour as preFourteenSumMeanMedianHour,

abs(t5.preFourteenMeanHour-t5.preFourteenMedianHour) as preFourteenDiffMeanMedianHour,

t5.preFourteenMinHour,

t5.preFourteenMaxHour,

t5.preFourteenMaxHour-t5.preFourteenMinHour as preFourteenDiffMaxMinHour

from

(

select

wifi\_ap\_tag,

hourTimeUsed,

avg(passenger\_count) as preDayBefMeanHour,

median(passenger\_count) as preDayBefMedianHour,

stddev(passenger\_count) as preDayBefStdHour

from

(

select wifi\_ap\_tag,hourTimeUsed,passenger\_count

from AJ\_wifi\_ap\_slice10min

where dateTimeUsed in ("11-10")

)tb1

group by wifi\_ap\_tag, hourTimeUsed

)t1

left outer join

(

select

wifi\_ap\_tag,

hourTimeUsed,

avg(passenger\_count) as preThreeMeanHour,

median(passenger\_count) as preThreeMedianHour,

stddev(passenger\_count) as preThreeStdHour

from

(

select wifi\_ap\_tag,hourTimeUsed,passenger\_count

from AJ\_wifi\_ap\_slice10min

where dateTimeUsed in ("11-08","11-09","11-10")

)tb2

group by wifi\_ap\_tag, hourTimeUsed

)t2

on t1.wifi\_ap\_tag = t2.wifi\_ap\_tag and t1.hourTimeUsed = t2.hourTimeUsed

left outer join

(

select

wifi\_ap\_tag,

hourTimeUsed,

avg(passenger\_count) as preFiveMeanHour,

median(passenger\_count) as preFiveMedianHour,

stddev(passenger\_count) as preFiveStdHour,

min(passenger\_count) as preFiveMinHour,

max(passenger\_count) as preFiveMaxHour

from

(

select wifi\_ap\_tag,hourTimeUsed,passenger\_count

from AJ\_wifi\_ap\_slice10min

where dateTimeUsed >= "11-06" and dateTimeUsed <="11-10"

)tb3

group by wifi\_ap\_tag, hourTimeUsed

)t3

on t1.wifi\_ap\_tag = t3.wifi\_ap\_tag and t1.hourTimeUsed = t3.hourTimeUsed

left outer join

(

select

wifi\_ap\_tag,

hourTimeUsed,

avg(passenger\_count) as preSevenMeanHour,

median(passenger\_count) as preSevenMedianHour,

stddev(passenger\_count) as preSevenStdHour,

min(passenger\_count) as preSevenMinHour,

max(passenger\_count) as preSevenMaxHour

from

(

select wifi\_ap\_tag,hourTimeUsed,passenger\_count

from AJ\_wifi\_ap\_slice10min

where dateTimeUsed >= "11-04" and dateTimeUsed <="11-10"

)tb4

group by wifi\_ap\_tag, hourTimeUsed

)t4

on t1.wifi\_ap\_tag = t4.wifi\_ap\_tag and t1.hourTimeUsed = t4.hourTimeUsed

left outer join

(

select

wifi\_ap\_tag,

hourTimeUsed,

avg(passenger\_count) as preFourteenMeanHour,

median(passenger\_count) as preFourteenMedianHour,

stddev(passenger\_count) as preFourteenStdHour,

min(passenger\_count) as preFourteenMinHour,

max(passenger\_count) as preFourteenMaxHour

from

(

select wifi\_ap\_tag,hourTimeUsed,passenger\_count

from AJ\_wifi\_ap\_slice10min

where dateTimeUsed >= "10-28" and dateTimeUsed <="11-10"

)tb5

group by wifi\_ap\_tag, hourTimeUsed

)t5

on t1.wifi\_ap\_tag = t5.wifi\_ap\_tag and t1.hourTimeUsed = t5.hourTimeUsed;

------------------------------------------------------------------------------------------

--按区域（wifi前5个字段）划分近1,3,5,7,14天按照“小时”的均值，中位数，标准差；

drop table if exists AJ\_FeatureUsedWithwifi\_11\_11;

create table if not exists AJ\_FeatureUsedWithwifi\_11\_11

as select

t1.hourTimeUsed,

t1.wifi\_ap\_tagObserve,

t1.preDayBefMeanwifi,

t1.preDayBefMedianwifi,

t1.preDayBefMeanwifi+t1.preDayBefMedianwifi as preDayBefSumMeanMedianwifi,

abs(t1.preDayBefMeanwifi-t1.preDayBefMedianwifi) as preDayBefDiffMeanMedianwifi,

t2.preThreeMeanwifi,

t2.preThreeMedianwifi,

t2.preThreeStdwifi,

t2.preThreeMeanwifi+t2.preThreeMedianwifi as preThreeSumMeanMedianwifi,

abs(t2.preThreeMeanwifi-t2.preThreeMedianwifi) as preThreeDiffMeanMedianwifi,

t3.preFiveMeanwifi,

t3.preFiveMedianwifi,

t3.preFiveStdwifi,

t3.preFiveMeanwifi+t3.preFiveMedianwifi as preFiveSumMeanMedianwifi,

abs(t3.preFiveMeanwifi-t3.preFiveMedianwifi) as preFiveDiffMeanMedianwifi,

t3.preFiveMinwifi,

t3.preFiveMaxwifi,

t3.preFiveMaxwifi-t3.preFiveMinwifi as preFiveDiffMaxMinwifi,

t4.preSevenMeanwifi,

t4.preSevenMedianwifi,

t4.preSevenStdwifi,

t4.preSevenMeanwifi+t4.preSevenMedianwifi as preSevenSumMeanMedianwifi,

abs(t4.preSevenMeanwifi-t4.preSevenMedianwifi) as preSevenDiffMeanMedianwifi,

t4.preSevenMinwifi,

t4.preSevenMaxwifi,

t4.preSevenMaxwifi-t4.preSevenMinwifi as preSevenDiffMaxMinwifi,

t5.preFourteenMeanwifi,

t5.preFourteenMedianwifi,

t5.preFourteenStdwifi,

t5.preFourteenMeanwifi+t5.preFourteenMedianwifi as preFourteenSumMeanMedianwifi,

abs(t5.preFourteenMeanwifi-t5.preFourteenMedianwifi) as preFourteenDiffMeanMedianwifi,

t5.preFourteenMinwifi,

t5.preFourteenMaxwifi,

t5.preFourteenMaxwifi-t5.preFourteenMinwifi as preFourteenDiffMaxMinwifi

from

(

select

hourTimeUsed,

wifi\_ap\_tagObserve,

avg(passenger\_count) as preDayBefMeanwifi,

median(passenger\_count) as preDayBefMedianwifi,

stddev(passenger\_count) as preDayBefStdwifi

from

(

select hourTimeUsed,passenger\_count,

substr(wifi\_ap\_tag, 1, 5) as wifi\_ap\_tagObserve

from AJ\_wifi\_ap\_slice10min

where dateTimeUsed in ("11-10")

)tb1

group by wifi\_ap\_tagObserve,hourTimeUsed

)t1

left outer join

(

select

hourTimeUsed,

wifi\_ap\_tagObserve,

avg(passenger\_count) as preThreeMeanwifi,

median(passenger\_count) as preThreeMedianwifi,

stddev(passenger\_count) as preThreeStdwifi

from

(

select hourTimeUsed,passenger\_count,

substr(wifi\_ap\_tag, 1, 5) as wifi\_ap\_tagObserve

from AJ\_wifi\_ap\_slice10min

where dateTimeUsed in ("11-08","11-09","11-10")

)tb2

group by wifi\_ap\_tagObserve,hourTimeUsed

)t2

on t1.hourTimeUsed = t2.hourTimeUsed and t1.wifi\_ap\_tagObserve = t2.wifi\_ap\_tagObserve

left outer join

(

select

hourTimeUsed,

wifi\_ap\_tagObserve,

avg(passenger\_count) as preFiveMeanwifi,

median(passenger\_count) as preFiveMedianwifi,

stddev(passenger\_count) as preFiveStdwifi,

min(passenger\_count) as preFiveMinwifi,

max(passenger\_count) as preFiveMaxwifi

from

(

select hourTimeUsed,passenger\_count,

substr(wifi\_ap\_tag, 1, 5) as wifi\_ap\_tagObserve

from AJ\_wifi\_ap\_slice10min

where dateTimeUsed >= "11-06" and dateTimeUsed <="11-10"

)tb3

group by wifi\_ap\_tagObserve,hourTimeUsed

)t3

on t1.hourTimeUsed = t3.hourTimeUsed and t1.wifi\_ap\_tagObserve = t3.wifi\_ap\_tagObserve

left outer join

(

select

hourTimeUsed,

wifi\_ap\_tagObserve,

avg(passenger\_count) as preSevenMeanwifi,

median(passenger\_count) as preSevenMedianwifi,

stddev(passenger\_count) as preSevenStdwifi,

min(passenger\_count) as preSevenMinwifi,

max(passenger\_count) as preSevenMaxwifi

from

(

select hourTimeUsed,passenger\_count,

substr(wifi\_ap\_tag, 1, 5) as wifi\_ap\_tagObserve

from AJ\_wifi\_ap\_slice10min

where dateTimeUsed >= "11-04" and dateTimeUsed <="11-10"

)tb4

group by wifi\_ap\_tagObserve,hourTimeUsed

)t4

on t1.hourTimeUsed = t4.hourTimeUsed and t1.wifi\_ap\_tagObserve = t4.wifi\_ap\_tagObserve

left outer join

(

select

hourTimeUsed,

wifi\_ap\_tagObserve,

avg(passenger\_count) as preFourteenMeanwifi,

median(passenger\_count) as preFourteenMedianwifi,

stddev(passenger\_count) as preFourteenStdwifi,

min(passenger\_count) as preFourteenMinwifi,

max(passenger\_count) as preFourteenMaxwifi

from

(

select hourTimeUsed,passenger\_count,

substr(wifi\_ap\_tag, 1, 5) as wifi\_ap\_tagObserve

from AJ\_wifi\_ap\_slice10min

where dateTimeUsed >= "10-28" and dateTimeUsed <="11-10"

)tb5

group by wifi\_ap\_tagObserve,hourTimeUsed

)t5

on t1.hourTimeUsed = t5.hourTimeUsed and t1.wifi\_ap\_tagObserve = t5.wifi\_ap\_tagObserve;

------------------------------------------------------------------------------------------

--按区域（wifi前5个字段）划分近1,3,5,7,14天按照“小时-分”的均值，中位数，标准差；

drop table if exists AJ\_FeatureUsedWithSectwifi\_11\_11;

create table if not exists AJ\_FeatureUsedWithSectwifi\_11\_11

as select

t1.hourTimeUsed,

t1.sect,

t1.wifi\_ap\_tagObserve,

t1.preDayBefMeanSectwifi,

t1.preDayBefMedianSectwifi,

t1.preDayBefMeanSectwifi+t1.preDayBefMedianSectwifi as preDayBefSumMeanMedianSectwifi,

abs(t1.preDayBefMeanSectwifi-t1.preDayBefMedianSectwifi) as preDayBefDiffMeanMedianSectwifi,

t2.preThreeMeanSectwifi,

t2.preThreeMedianSectwifi,

t2.preThreeStdSectwifi,

t2.preThreeMeanSectwifi+t2.preThreeMedianSectwifi as preThreeSumMeanMedianSectwifi,

abs(t2.preThreeMeanSectwifi-t2.preThreeMedianSectwifi) as preThreeDiffMeanMedianSectwifi,

t3.preFiveMeanSectwifi,

t3.preFiveMedianSectwifi,

t3.preFiveStdSectwifi,

t3.preFiveMeanSectwifi+t3.preFiveMedianSectwifi as preFiveSumMeanMedianSectwifi,

abs(t3.preFiveMeanSectwifi-t3.preFiveMedianSectwifi) as preFiveDiffMeanMedianSectwifi,

t3.preFiveMinSectwifi,

t3.preFiveMaxSectwifi,

t3.preFiveMaxSectwifi-t3.preFiveMinSectwifi as preFiveDiffMaxMinSectwifi,

t4.preSevenMeanSectwifi,

t4.preSevenMedianSectwifi,

t4.preSevenStdSectwifi,

t4.preSevenMeanSectwifi+t4.preSevenMedianSectwifi as preSevenSumMeanMedianSectwifi,

abs(t4.preSevenMeanSectwifi-t4.preSevenMedianSectwifi) as preSevenDiffMeanMedianSectwifi,

t4.preSevenMinSectwifi,

t4.preSevenMaxSectwifi,

t4.preSevenMaxSectwifi-t4.preSevenMinSectwifi as preSevenDiffMaxMinSectwifi,

t5.preFourteenMeanSectwifi,

t5.preFourteenMedianSectwifi,

t5.preFourteenStdSectwifi,

t5.preFourteenMeanSectwifi+t5.preFourteenMedianSectwifi as preFourteenSumMeanMedianSectwifi,

abs(t5.preFourteenMeanSectwifi-t5.preFourteenMedianSectwifi) as preFourteenDiffMeanMedianSectwifi,

t5.preFourteenMinSectwifi,

t5.preFourteenMaxSectwifi,

t5.preFourteenMaxSectwifi-t5.preFourteenMinSectwifi as preFourteenDiffMaxMinSectwifi

from

(

select

hourTimeUsed,

wifi\_ap\_tagObserve,

sect,

avg(passenger\_count) as preDayBefMeanSectwifi,

median(passenger\_count) as preDayBefMedianSectwifi,

stddev(passenger\_count) as preDayBefStdSectwifi

from

(

select hourTimeUsed,passenger\_count,sect,

substr(wifi\_ap\_tag, 1, 5) as wifi\_ap\_tagObserve

from AJ\_wifi\_ap\_slice10min

where dateTimeUsed in ("11-10")

)tb1

group by wifi\_ap\_tagObserve,sect,hourTimeUsed

)t1

left outer join

(

select

hourTimeUsed,

wifi\_ap\_tagObserve,

sect,

avg(passenger\_count) as preThreeMeanSectwifi,

median(passenger\_count) as preThreeMedianSectwifi,

stddev(passenger\_count) as preThreeStdSectwifi

from

(

select hourTimeUsed,passenger\_count,sect,

substr(wifi\_ap\_tag, 1, 5) as wifi\_ap\_tagObserve

from AJ\_wifi\_ap\_slice10min

where dateTimeUsed in ("11-08","11-09","11-10")

)tb2

group by wifi\_ap\_tagObserve,sect,hourTimeUsed

)t2

on t1.hourTimeUsed = t2.hourTimeUsed and t1.wifi\_ap\_tagObserve = t2.wifi\_ap\_tagObserve and t1.sect = t2.sect

left outer join

(

select

hourTimeUsed,

wifi\_ap\_tagObserve,

sect,

avg(passenger\_count) as preFiveMeanSectwifi,

median(passenger\_count) as preFiveMedianSectwifi,

stddev(passenger\_count) as preFiveStdSectwifi,

min(passenger\_count) as preFiveMinSectwifi,

max(passenger\_count) as preFiveMaxSectwifi

from

(

select hourTimeUsed,passenger\_count,sect,

substr(wifi\_ap\_tag, 1, 5) as wifi\_ap\_tagObserve

from AJ\_wifi\_ap\_slice10min

where dateTimeUsed >= "11-06" and dateTimeUsed <="11-10"

)tb3

group by wifi\_ap\_tagObserve,sect,hourTimeUsed

)t3

on t1.hourTimeUsed = t3.hourTimeUsed and t1.wifi\_ap\_tagObserve = t3.wifi\_ap\_tagObserve and t1.sect = t3.sect

left outer join

(

select

hourTimeUsed,

wifi\_ap\_tagObserve,

sect,

avg(passenger\_count) as preSevenMeanSectwifi,

median(passenger\_count) as preSevenMedianSectwifi,

stddev(passenger\_count) as preSevenStdSectwifi,

min(passenger\_count) as preSevenMinSectwifi,

max(passenger\_count) as preSevenMaxSectwifi

from

(

select hourTimeUsed,passenger\_count,sect,

substr(wifi\_ap\_tag, 1, 5) as wifi\_ap\_tagObserve

from AJ\_wifi\_ap\_slice10min

where dateTimeUsed >= "11-04" and dateTimeUsed <="11-10"

)tb4

group by wifi\_ap\_tagObserve,sect,hourTimeUsed

)t4

on t1.hourTimeUsed = t4.hourTimeUsed and t1.wifi\_ap\_tagObserve = t4.wifi\_ap\_tagObserve and t1.sect = t4.sect

left outer join

(

select

hourTimeUsed,

wifi\_ap\_tagObserve,

sect,

avg(passenger\_count) as preFourteenMeanSectwifi,

median(passenger\_count) as preFourteenMedianSectwifi,

stddev(passenger\_count) as preFourteenStdSectwifi,

min(passenger\_count) as preFourteenMinSectwifi,

max(passenger\_count) as preFourteenMaxSectwifi

from

(

select hourTimeUsed,passenger\_count,sect,

substr(wifi\_ap\_tag, 1, 5) as wifi\_ap\_tagObserve

from AJ\_wifi\_ap\_slice10min

where dateTimeUsed >= "10-28" and dateTimeUsed <="11-10"

)tb5

group by wifi\_ap\_tagObserve,sect,hourTimeUsed

)t5

on t1.hourTimeUsed = t5.hourTimeUsed and t1.wifi\_ap\_tagObserve = t5.wifi\_ap\_tagObserve and t1.sect = t5.sect;

-----------------------------------------------------------------------------------------

--如何体现星期特征~

drop table if exists AJ\_FeatureUsedWithWeekday\_11\_11;

create table if not exists AJ\_FeatureUsedWithWeekday\_11\_11 as

select

wifi\_ap\_tag,

sect,

hourTimeUsed,

weekdaytime,

preWeekdayMean,

preWeekdayMedian,

preWeekdayStd,

preWeekdayMin,

preWeekdayMax,

preWeekdayMean+preWeekdayMedian as preWeekdaySumMeanMedian,

abs(preWeekdayMean-preWeekdayMedian) as preWeekdayDiffMeanMedian,

preWeekdayMax-preWeekdayMin as preWeekdayDiffMaxMin

from

(

select

wifi\_ap\_tag,

sect,

hourTimeUsed,

weekdaytime,

avg(passenger\_count) as preWeekdayMean,

median(passenger\_count) as preWeekdayMedian,

stddev(passenger\_count) as preWeekdayStd,

min(passenger\_count) as preWeekdayMin,

max(passenger\_count) as preWeekdayMax

from AJ\_wifi\_ap\_slice10min

where dateTimeUsed < "11-11"

group by wifi\_ap\_tag, sect, hourTimeUsed,weekdaytime

)tb;

------------------------------------------------------------------------------------------

--如何体现间隔特征~

drop table if exists AJ\_FeatureUsedWithInterval\_11\_11;

create table if not exists AJ\_FeatureUsedWithInterval\_11\_11

as select

wifi\_ap\_tag,

sect,

hourTimeUsed,

dateTimeUsed,

preIntervalMean,

preIntervalMedian,

preIntervalStd,

preIntervalMax,

preIntervalMin,

preIntervalMean+preIntervalMedian as preIntervalSumMeanMedian,

abs(preIntervalMean-preIntervalMedian) as preIntervalDiffMeanMedian,

preIntervalMax-preIntervalMin as preIntervalDiffMaxMin

from

(

select

wifi\_ap\_tag,

sect,

hourTimeUsed,

'11-12' as dateTimeUsed,

avg(passenger\_count) as preIntervalMean,

median(passenger\_count) as preIntervalMedian,

stddev(passenger\_count) as preIntervalStd,

max(passenger\_count) as preIntervalMax,

min(passenger\_count) as preIntervalMin

from AJ\_wifi\_ap\_slice10min

where dateTimeUsed in("11-10","11-08","11-06","11-04","11-02")

group by wifi\_ap\_tag, sect, hourTimeUsed

union all

select

wifi\_ap\_tag,

sect,

hourTimeUsed,

'11-11' as dateTimeUsed,

avg(passenger\_count) as preIntervalMean,

median(passenger\_count) as preIntervalMedian,

stddev(passenger\_count) as preIntervalStd,

max(passenger\_count) as preIntervalMax,

min(passenger\_count) as preIntervalMin

from AJ\_wifi\_ap\_slice10min

where dateTimeUsed in("11-09","11-07","11-05","11-03","11-01")

group by wifi\_ap\_tag, sect, hourTimeUsed

)tb;

-------------------------------------------------------------

-------------------------------------------------------------

-------------------------------------------------------------

-- 特征融合

-------------------------------------------------------------

-------------------------------------------------------------

-------------------------------------------------------------

drop table if exists AJ\_onlineUsedFormat;

create table AJ\_onlineUsedFormat as

select

wifi\_ap\_tag,

sect,

hourTimeUsed,

hourMinute,

substr(wifi\_ap\_tag, 1, 5) as wifi\_ap\_tagObserve,

concat(substr(slice10min, 1, 8),

cast((cast(substr(slice10min, 9, 2) as bigint)+2) as string),

substr(slice10min, 11)) as slice10min,

concat(substr(slice10min, 6, 3),

cast((cast(substr(slice10min, 9, 2) as bigint)+2) as string)) as dateTimeUsed

from AJ\_wifi\_ap\_slice10min

where slice10min >= "2016-11-09-00-0" and slice10min <= "2016-11-10-23-5";

drop table if exists AJ\_onlineTestUsed;

create table if not exists AJ\_onlineTestUsed as

select wifi\_ap\_tag, sect, dateTimeUsed, hourTimeUsed, hourMinute, wifi\_ap\_tagObserve, slice10min,

case when dateTimeUsed in ("11-11") then 4

when dateTimeUsed in ("11-12") then 5

end as weekDayTime

from AJ\_onlineUsedFormat;

drop table if exists AJ\_onlineTestFeatureCombineUsed;

create table AJ\_onlineTestFeatureCombineUsed as

select

t1.wifi\_ap\_tag,

t1.wifi\_ap\_tagObserve,

t1.slice10min,

t1.sect,

t1.dateTimeUsed,

t1.hourTimeUsed,

t1.hourMinute,

t1.weekDayTime,

cast(t1.hourTimeUsed as bigint) as hourInt,

cast(t1.sect as bigint) as sectInt,

cast(substr(t1.dateTimeUsed,1,2) as bigint) as monthInt,

cast(substr(t1.dateTimeUsed,4,2) as bigint) as dayInt,

case when t1.dateTimeUsed in ("09-15","09-16","09-17","10-01","10-02","10-03","10-04","10-05","10-06","10-07") then 1 else 0

end as isHoliday,

case when t1.dateTimeUsed in ("09-14", "09-29", "09-30") then 1 else 0

end as isBefHoliday,

case when t1.dateTimeUsed in ("09-18", "10-08", "10-09") then 1 else 0

end as isAfterHoliday,

t2.preDayBefMean,

t2.preDayBefMedian,

t2.preDayBefSumMeanMedian,

t2.preDayBefDiffMeanMedian,

t2.preThreeMean,

t2.preThreeMedian,

t2.preThreeStd,

t2.preThreeSumMeanMedian,

t2.preThreeDiffMeanMedian,

t2.preFiveMean,

t2.preFiveMedian,

t2.preFiveStd,

t2.preFiveSumMeanMedian,

t2.preFiveDiffMeanMedian,

t2.preFiveMin,

t2.preFiveMax,

t2.preFiveDiffMaxMin,

t2.preSevenMean,

t2.preSevenMedian,

t2.preSevenSumMeanMedian,

t2.preSevenDiffMeanMedian,

t2.preSevenMin,

t2.preSevenMax,

t2.preSevenDiffMaxMin,

t2.preFourteenMean,

t2.preFourteenMedian,

t2.preFourteenStd,

t2.preFourteenSumMeanMedian,

t2.preFourteenDiffMeanMedian,

t2.preFourteenMin,

t2.preFourteenMax,

t2.preFourteenDiffMaxMin,

t3.preDayBefMeanHour,

t3.preDayBefMedianHour,

t3.preDayBefSumMeanMedianHour,

t3.preDayBefDiffMeanMedianHour,

t3.preThreeMeanHour,

t3.preThreeMedianHour,

t3.preThreeStdHour,

t3.preThreeSumMeanMedianHour,

t3.preThreeDiffMeanMedianHour,

t3.preFiveMeanHour,

t3.preFiveMedianHour,

t3.preFiveStdHour,

t3.preFiveSumMeanMedianHour,

t3.preFiveDiffMeanMedianHour,

t3.preFiveMinHour,

t3.preFiveMaxHour,

t3.preFiveDiffMaxMinHour,

t3.preSevenMeanHour,

t3.preSevenMedianHour,

t3.preSevenSumMeanMedianHour,

t3.preSevenDiffMeanMedianHour,

t3.preSevenMinHour,

t3.preSevenMaxHour,

t3.preSevenDiffMaxMinHour,

t3.preFourteenMeanHour,

t3.preFourteenMedianHour,

t3.preFourteenStdHour,

t3.preFourteenSumMeanMedianHour,

t3.preFourteenDiffMeanMedianHour,

t3.preFourteenMinHour,

t3.preFourteenMaxHour,

t3.preFourteenDiffMaxMinHour,

t4.preDayBefMeanSectwifi,

t4.preDayBefMedianSectwifi,

t4.preDayBefSumMeanMedianSectwifi,

t4.preDayBefDiffMeanMedianSectwifi,

t4.preThreeMeanSectwifi,

t4.preThreeMedianSectwifi,

t4.preThreeStdSectwifi,

t4.preThreeSumMeanMedianSectwifi,

t4.preThreeDiffMeanMedianSectwifi,

t4.preFiveMeanSectwifi,

t4.preFiveMedianSectwifi,

t4.preFiveStdSectwifi,

t4.preFiveSumMeanMedianSectwifi,

t4.preFiveDiffMeanMedianSectwifi,

t4.preFiveMinSectwifi,

t4.preFiveMaxSectwifi,

t4.preFiveDiffMaxMinSectwifi,

t4.preSevenMeanSectwifi,

t4.preSevenMedianSectwifi,

t4.preSevenSumMeanMedianSectwifi,

t4.preSevenDiffMeanMedianSectwifi,

t4.preSevenMinSectwifi,

t4.preSevenMaxSectwifi,

t4.preSevenDiffMaxMinSectwifi,

t4.preFourteenMeanSectwifi,

t4.preFourteenMedianSectwifi,

t4.preFourteenStdSectwifi,

t4.preFourteenSumMeanMedianSectwifi,

t4.preFourteenDiffMeanMedianSectwifi,

t4.preFourteenMinSectwifi,

t4.preFourteenMaxSectwifi,

t4.preFourteenDiffMaxMinSectwifi,

t5.preDayBefMeanwifi,

t5.preDayBefMedianwifi,

t5.preDayBefSumMeanMedianwifi,

t5.preDayBefDiffMeanMedianwifi,

t5.preThreeMeanwifi,

t5.preThreeMedianwifi,

t5.preThreeStdwifi,

t5.preThreeSumMeanMedianwifi,

t5.preThreeDiffMeanMedianwifi,

t5.preFiveMeanwifi,

t5.preFiveMedianwifi,

t5.preFiveStdwifi,

t5.preFiveSumMeanMedianwifi,

t5.preFiveDiffMeanMedianwifi,

t5.preFiveMinwifi,

t5.preFiveMaxwifi,

t5.preFiveDiffMaxMinwifi,

t5.preSevenMeanwifi,

t5.preSevenMedianwifi,

t5.preSevenSumMeanMedianwifi,

t5.preSevenDiffMeanMedianwifi,

t5.preSevenMinwifi,

t5.preSevenMaxwifi,

t5.preSevenDiffMaxMinwifi,

t5.preFourteenMeanwifi,

t5.preFourteenMedianwifi,

t5.preFourteenStdwifi,

t5.preFourteenSumMeanMedianwifi,

t5.preFourteenDiffMeanMedianwifi,

t5.preFourteenMinwifi,

t5.preFourteenMaxwifi,

t5.preFourteenDiffMaxMinwifi,

m1.oh\_wifi\_ap\_tag,

m2.oh\_wifi\_ap\_tagObserve,

m3.oh\_hourMinute,

t6.preWeekdayMean,

t6.preWeekdayMedian,

t6.preWeekdayStd,

t6.preWeekdayMin,

t6.preWeekdayMax,

t6.preWeekdaySumMeanMedian,

t6.preWeekdayDiffMeanMedian,

t6.preWeekdayDiffMaxMin,

t7.preIntervalMean,

t7.preIntervalMedian,

t7.preIntervalStd,

t7.preIntervalMax,

t7.preIntervalMin,

t7.preIntervalSumMeanMedian,

t7.preIntervalDiffMeanMedian,

t7.preIntervalDiffMaxMin

from AJ\_onlineTestUsed t1

left outer join AJ\_FeatureUsedWithSect\_11\_11 t2

on t1.wifi\_ap\_tag=t2.wifi\_ap\_tag and t1.sect = t2.sect and t1.hourTimeUsed = t2.hourTimeUsed

left outer join AJ\_FeatureUsedWithHour\_11\_11 t3

on t1.wifi\_ap\_tag=t3.wifi\_ap\_tag and t1.hourTimeUsed = t3.hourTimeUsed

left outer join AJ\_FeatureUsedWithSectwifi\_11\_11 t4

on t1.wifi\_ap\_tagObserve=t4.wifi\_ap\_tagObserve and t1.sect = t4.sect and t1.hourTimeUsed = t4.hourTimeUsed

left outer join AJ\_FeatureUsedWithwifi\_11\_11 t5

on t1.wifi\_ap\_tagObserve=t5.wifi\_ap\_tagObserve and t1.hourTimeUsed = t5.hourTimeUsed

left outer join AJ\_wifiAPTagMap m1

on t1.wifi\_ap\_tag=m1.wifi\_ap\_tag

left outer join aj\_wifiAPTagObserveMap m2

on t1.wifi\_ap\_tagObserve=m2.wifi\_ap\_tagObserve

left outer join aj\_hourMinuteMap m3

on t1.hourMinute=m3.hourMinute

left outer join AJ\_FeatureUsedWithWeekday\_11\_11 t6

on t1.wifi\_ap\_tag=t6.wifi\_ap\_tag and t1.sect = t6.sect and t1.hourTimeUsed = t6.hourTimeUsed and t1.weekDayTime=t6.weekDayTime

left outer join AJ\_FeatureUsedWithInterval\_11\_11 t7

on t1.wifi\_ap\_tag=t7.wifi\_ap\_tag and t1.sect = t7.sect and t1.hourTimeUsed = t7.hourTimeUsed and t1.dateTimeUsed=t7.dateTimeUsed;

4.2登机区域模型

特征提取，以线上测试集为例：

drop table if exists AJ\_onlineTestArea\_11\_11\_12;

create table AJ\_onlineTestArea\_11\_11\_12 as

select

t1.wifi\_ap\_tag,

t1.slice10min,

t1.oh\_wifi\_ap\_tag,

t1.oh\_hourMinute,

t1.preSevenMean,

t21.flights\_num\_aft\_30,

t22.flights\_num\_pre\_30,

t23.flights\_num\_pre\_60,

t24.flights\_num\_pre\_90,

t25.flights\_num\_pre\_120,

t26.flights\_num\_pre\_150,

t27.flights\_num\_pre\_180,

t31.passenger\_num\_aft\_30,

t32.passenger\_num\_pre\_30,

t33.passenger\_num\_pre\_60,

t34.passenger\_num\_pre\_90,

t35.passenger\_num\_pre\_120,

t36.passenger\_num\_pre\_150,

t37.passenger\_num\_pre\_180

from

(

select \* from AJ\_onlineTestFeatureCombineUsed

where wifi\_ap\_tag in (select wifi\_ap\_tag from lyd\_wifi\_bgate\_relation\_new)

)t1

left outer join

(

select wifi\_ap\_tag, slice10min,

sum(flight\_num) as flights\_num\_aft\_30

from

aj\_wifi\_bgate\_relation tb1

left outer join

aj\_airport\_gz\_flights\_num\_aft\_30 tb2

on tb1.bgate\_id=tb2.bgate\_id

group by wifi\_ap\_tag, slice10min

)t21

on t1.wifi\_ap\_tag=t21.wifi\_ap\_tag and t1.slice10min=t21.slice10min

left outer join

(

select wifi\_ap\_tag, slice10min,

sum(flight\_num) as flights\_num\_pre\_30

from

aj\_wifi\_bgate\_relation tb1

left outer join

aj\_airport\_gz\_flights\_num\_pre\_30 tb2

on tb1.bgate\_id=tb2.bgate\_id

group by wifi\_ap\_tag, slice10min

)t22

on t1.wifi\_ap\_tag=t22.wifi\_ap\_tag and t1.slice10min=t22.slice10min

left outer join

(

select wifi\_ap\_tag, slice10min,

sum(flight\_num) as flights\_num\_pre\_60

from

aj\_wifi\_bgate\_relation tb1

left outer join

aj\_airport\_gz\_flights\_num\_pre\_60 tb2

on tb1.bgate\_id=tb2.bgate\_id

group by wifi\_ap\_tag, slice10min

)t23

on t1.wifi\_ap\_tag=t23.wifi\_ap\_tag and t1.slice10min=t23.slice10min

left outer join

(

select wifi\_ap\_tag, slice10min,

sum(flight\_num) as flights\_num\_pre\_90

from

aj\_wifi\_bgate\_relation tb1

left outer join

aj\_airport\_gz\_flights\_num\_pre\_90 tb2

on tb1.bgate\_id=tb2.bgate\_id

group by wifi\_ap\_tag, slice10min

)t24

on t1.wifi\_ap\_tag=t24.wifi\_ap\_tag and t1.slice10min=t24.slice10min

left outer join

(

select wifi\_ap\_tag, slice10min,

sum(flight\_num) as flights\_num\_pre\_120

from

aj\_wifi\_bgate\_relation tb1

left outer join

aj\_airport\_gz\_flights\_num\_pre\_120 tb2

on tb1.bgate\_id=tb2.bgate\_id

group by wifi\_ap\_tag, slice10min

)t25

on t1.wifi\_ap\_tag=t25.wifi\_ap\_tag and t1.slice10min=t25.slice10min

left outer join

(

select wifi\_ap\_tag, slice10min,

sum(flight\_num) as flights\_num\_pre\_150

from

aj\_wifi\_bgate\_relation tb1

left outer join

aj\_airport\_gz\_flights\_num\_pre\_150 tb2

on tb1.bgate\_id=tb2.bgate\_id

group by wifi\_ap\_tag, slice10min

)t26

on t1.wifi\_ap\_tag=t26.wifi\_ap\_tag and t1.slice10min=t26.slice10min

left outer join

(

select wifi\_ap\_tag, slice10min,

sum(flight\_num) as flights\_num\_pre\_180

from

aj\_wifi\_bgate\_relation tb1

left outer join

aj\_airport\_gz\_flights\_num\_pre\_180 tb2

on tb1.bgate\_id=tb2.bgate\_id

group by wifi\_ap\_tag, slice10min

)t27

on t1.wifi\_ap\_tag=t27.wifi\_ap\_tag and t1.slice10min=t27.slice10min

left outer join

(

select wifi\_ap\_tag, slice10min,

sum(passenger\_num) as passenger\_num\_aft\_30

from

aj\_wifi\_bgate\_relation tb1

left outer join

aj\_airport\_gz\_passenger\_num\_aft\_30 tb2

on tb1.bgate\_id=tb2.bgate\_id

group by wifi\_ap\_tag, slice10min

)t31

on t1.wifi\_ap\_tag=t31.wifi\_ap\_tag and t1.slice10min=t31.slice10min

left outer join

(

select wifi\_ap\_tag, slice10min,

sum(passenger\_num) as passenger\_num\_pre\_30

from

aj\_wifi\_bgate\_relation tb1

left outer join

aj\_airport\_gz\_passenger\_num\_pre\_30 tb2

on tb1.bgate\_id=tb2.bgate\_id

group by wifi\_ap\_tag, slice10min

)t32

on t1.wifi\_ap\_tag=t32.wifi\_ap\_tag and t1.slice10min=t32.slice10min

left outer join

(

select wifi\_ap\_tag, slice10min,

sum(passenger\_num) as passenger\_num\_pre\_60

from

aj\_wifi\_bgate\_relation tb1

left outer join

aj\_airport\_gz\_passenger\_num\_pre\_60 tb2

on tb1.bgate\_id=tb2.bgate\_id

group by wifi\_ap\_tag, slice10min

)t33

on t1.wifi\_ap\_tag=t33.wifi\_ap\_tag and t1.slice10min=t33.slice10min

left outer join

(

select wifi\_ap\_tag, slice10min,

sum(passenger\_num) as passenger\_num\_pre\_90

from

aj\_wifi\_bgate\_relation tb1

left outer join

aj\_airport\_gz\_passenger\_num\_pre\_90 tb2

on tb1.bgate\_id=tb2.bgate\_id

group by wifi\_ap\_tag, slice10min

)t34

on t1.wifi\_ap\_tag=t34.wifi\_ap\_tag and t1.slice10min=t34.slice10min

left outer join

(

select wifi\_ap\_tag, slice10min,

sum(passenger\_num) as passenger\_num\_pre\_120

from

aj\_wifi\_bgate\_relation tb1

left outer join

aj\_airport\_gz\_passenger\_num\_pre\_120 tb2

on tb1.bgate\_id=tb2.bgate\_id

group by wifi\_ap\_tag, slice10min

)t35

on t1.wifi\_ap\_tag=t35.wifi\_ap\_tag and t1.slice10min=t35.slice10min

left outer join

(

select wifi\_ap\_tag, slice10min,

sum(passenger\_num) as passenger\_num\_pre\_150

from

aj\_wifi\_bgate\_relation tb1

left outer join

aj\_airport\_gz\_passenger\_num\_pre\_150 tb2

on tb1.bgate\_id=tb2.bgate\_id

group by wifi\_ap\_tag, slice10min

)t36

on t1.wifi\_ap\_tag=t36.wifi\_ap\_tag and t1.slice10min=t36.slice10min

left outer join

(

select wifi\_ap\_tag, slice10min,

sum(passenger\_num) as passenger\_num\_pre\_180

from

aj\_wifi\_bgate\_relation tb1

left outer join

aj\_airport\_gz\_passenger\_num\_pre\_180 tb2

on tb1.bgate\_id=tb2.bgate\_id

group by wifi\_ap\_tag, slice10min

)t37

on t1.wifi\_ap\_tag=t37.wifi\_ap\_tag and t1.slice10min=t37.slice10min;

select \* from AJ\_onlineTestArea\_11\_11\_12

order by wifi\_ap\_tag, slice10min limit 1000000;