

# 苏州信令数据质量分析报告

## 1 基础分析

### 1.1 用户数量

根据 2015 年 1 月 22 日 0 : 00 至 23 : 00 约一天的信令数据中，全天共采集到约 690 万名手机用户，共 2 亿 5 千万条信令记录。人均信令量约为 250 条/天。

一天用户量	一天信令总量	人均信令数
6908801 人	255120820 条	36 条/人·天

\*只统计目前获得的五类信令数据的样本

### 1.2 用户归属地分析

本次提供的信令数据中，还包含了用户的归属地。根据现有数据及用户归属地信息，分析结果如图 1 所示：

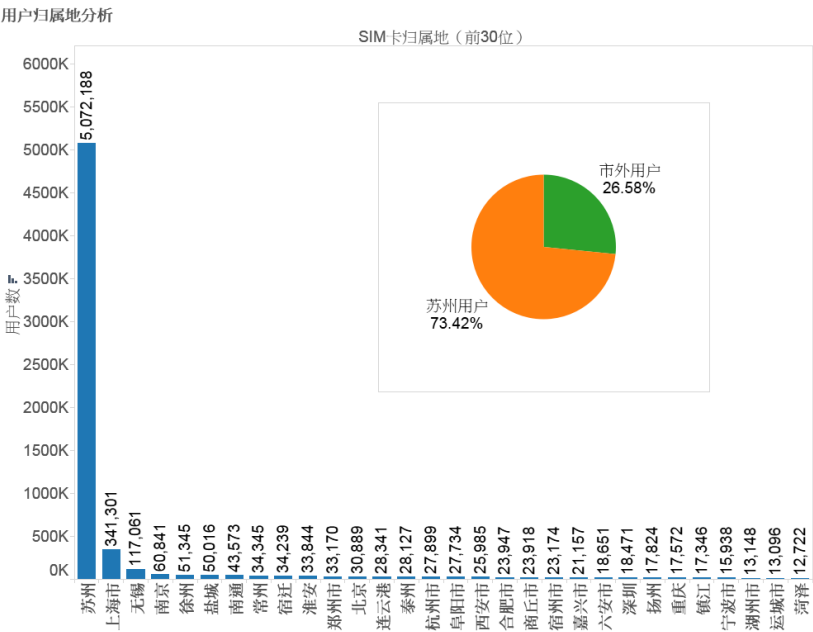


图 1：用户归属地分析

在一天所采集的数据中发现，苏州市外的手机用户占 25%以上，用户遍布全国，苏州市拥护强大的人口与客流吸引能力；信令数据中包括手机用户的归属地信息，可辅助进行用户交通与出行状况的深入分析。

## 2 基站分布情况分析

### 2.1 2G 与 3G 基站分布情况

本次所获取数据包含了 2G 数据与 TD 数据（即移动 3G 数据）。情况如下：

	基站数量	基站量占比	所采集信令数量	信令占比
2G	27,468 个	56%	147,525,551 条	60%
3G	21,194 个	44%	98,127,611 条	40%

根据统计分析，3G 基站以及所贡献的信令量，均约占总比 40%，是数据的重要组成部分。

### 2.2 各区域的基站情况

图 2 与图 3 分别是 2G、3G 基站在苏州市的空间分布情况。

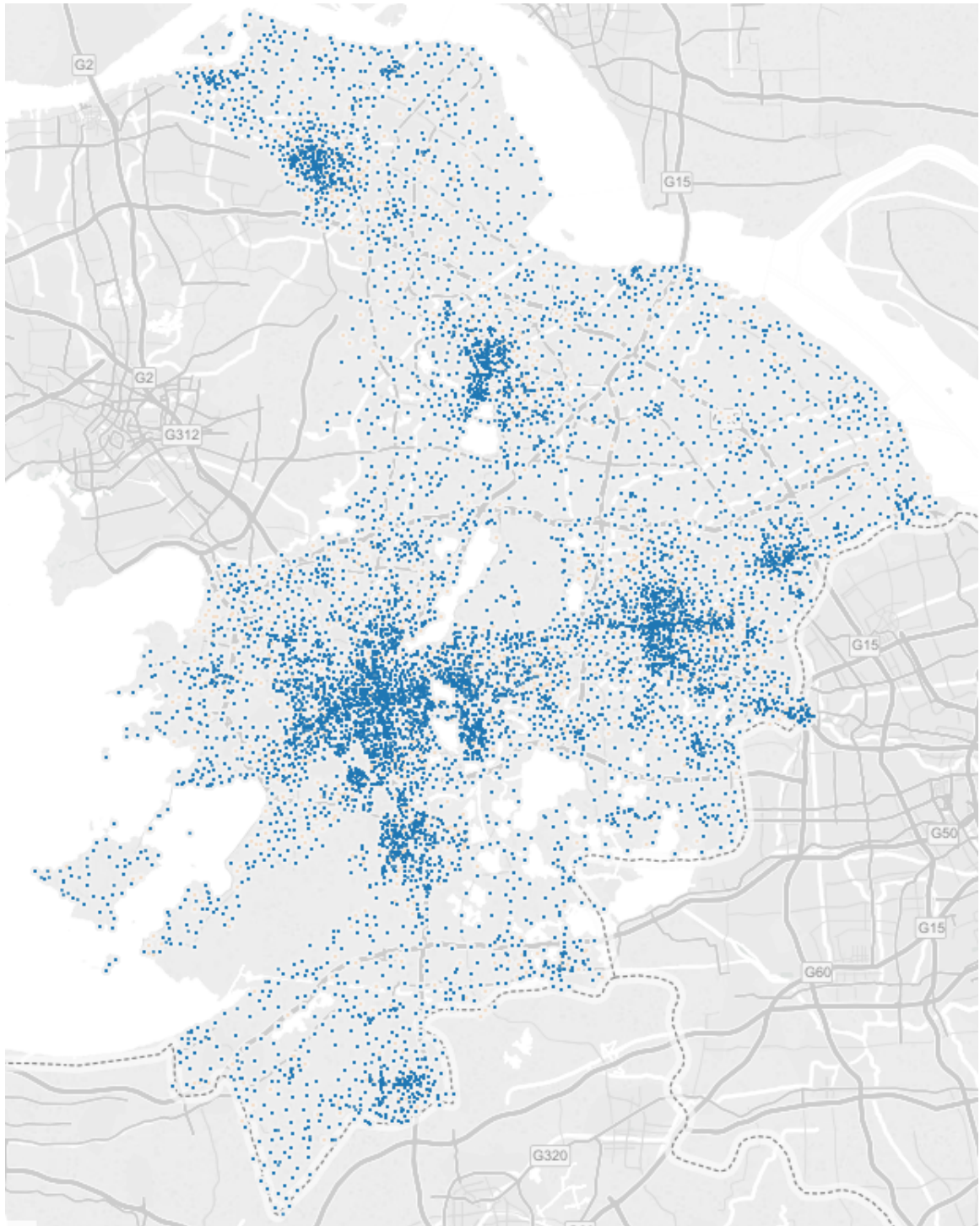


图 2：2G 基站分布情况

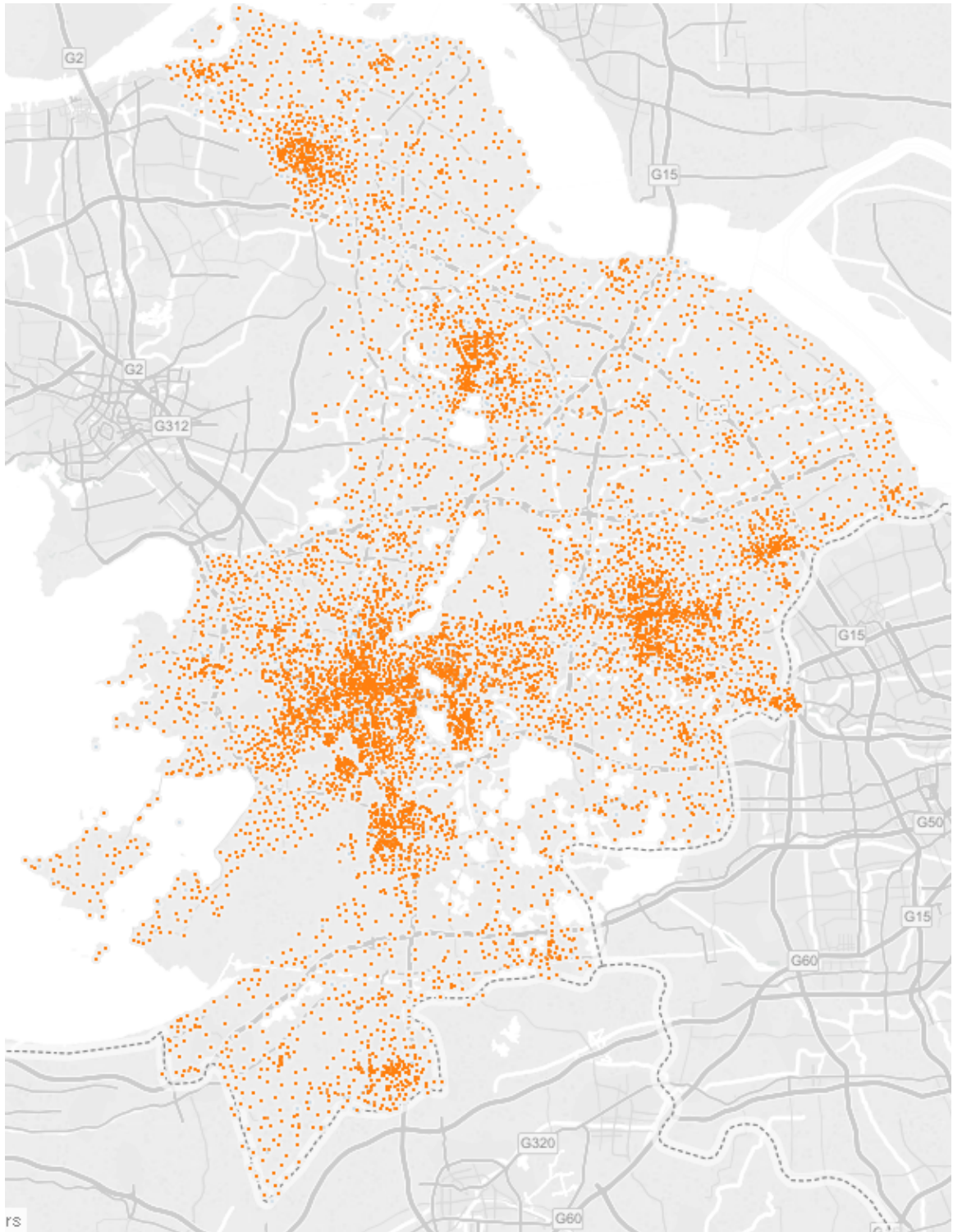


图 3 : 3G 基站分布情况

图 2 与图 3 反映了 2G 基站与 3G 基站的建设密度与位置分布基本一致。

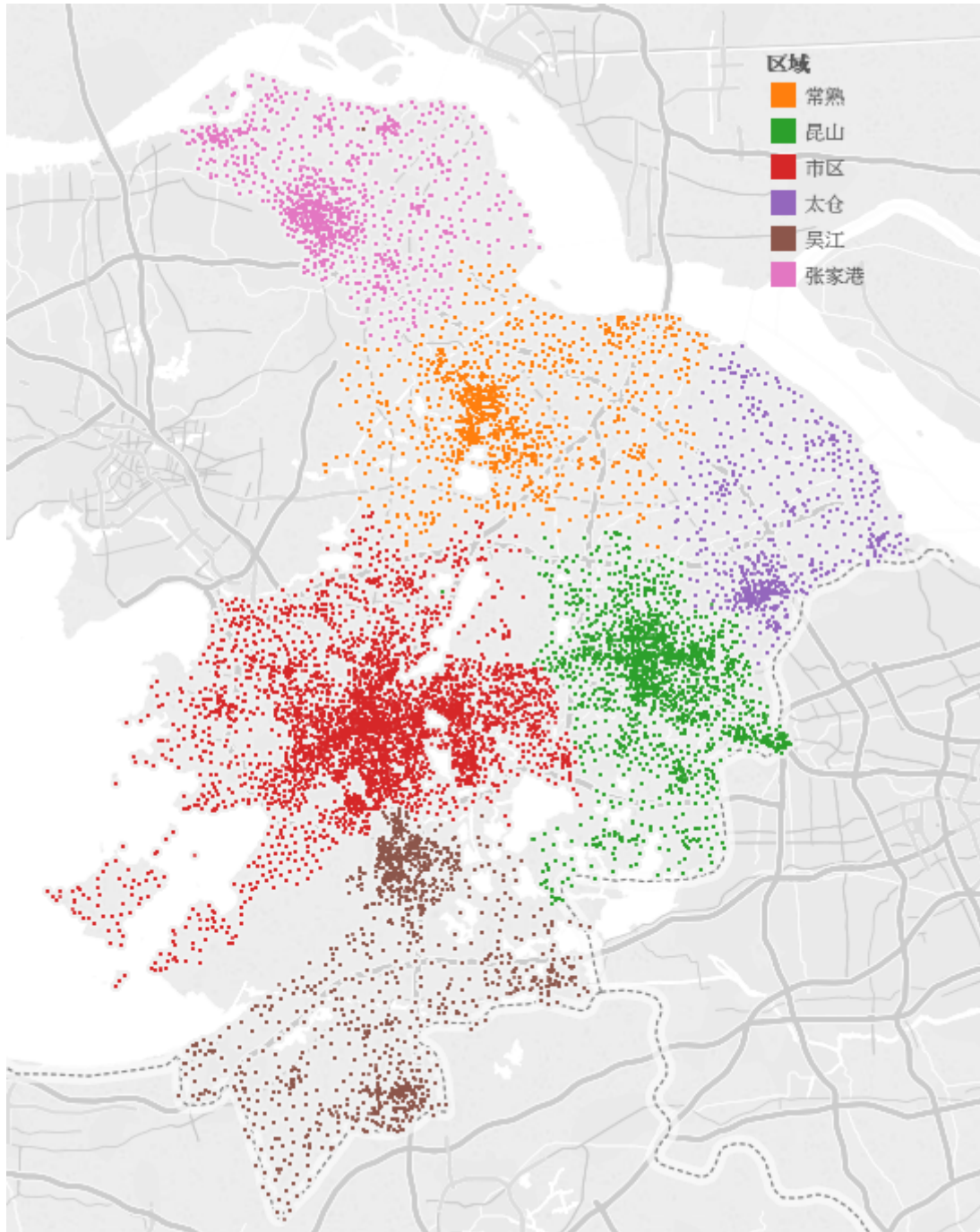


图 4：苏州各区域的基站分布情况（包含 2G 与 3G）

而从图 4 中可以看到，苏州市区与昆山的基站建设最为密集，与人口密度状况极为吻合。

### 3 信令情况分析

#### 3.1 各事件比例情况

根据目前所获取的五个信令数据文件，包括五类信令事件（正常位置更新、切换、voice、

SMS、detach) 进行了初步的统计，正常位置更新事件数量占五类事件数量的 66%，一天共收集 1.6 亿条记录。具体情况见下图 5 所示。

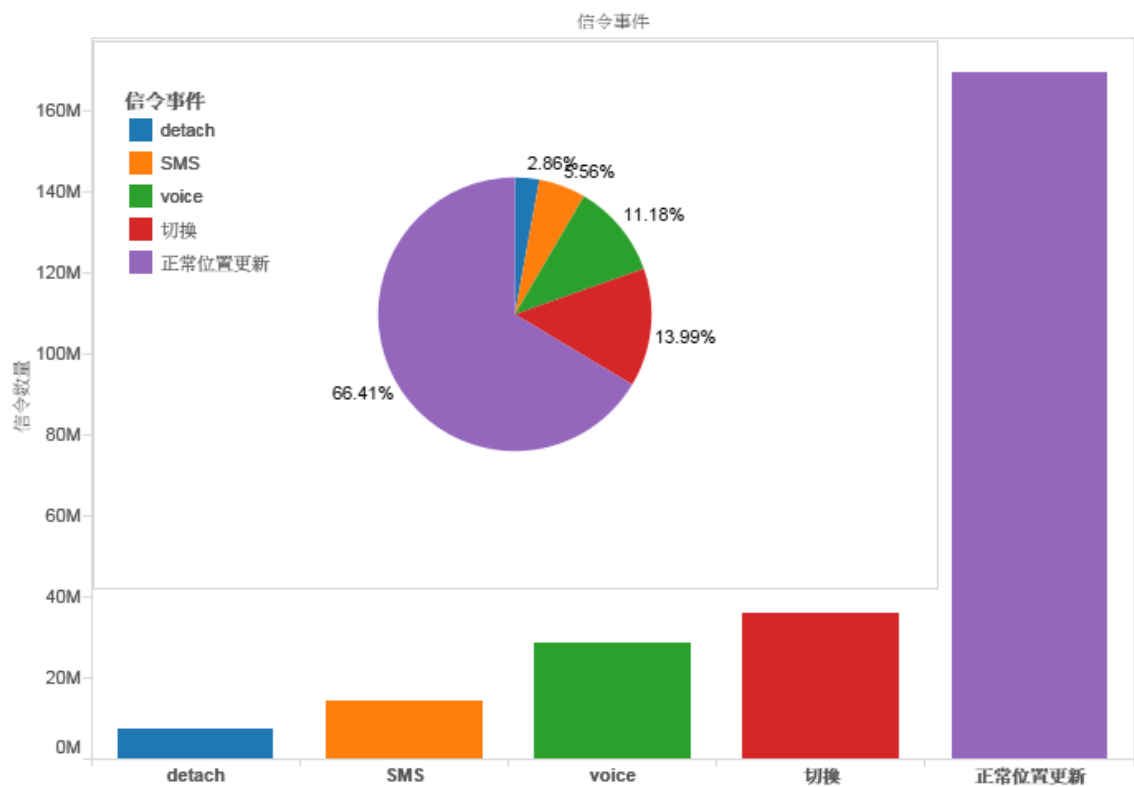


图 5：各信令事件比例情况

由于目前只获取五类事件的数据文件，在缺少全样本数据的情况下，暂时无法进行如周期性位置更新丢失率等更深入数据质量评估。因此，需要获得更全的数据，方可进行更深入的分析。

### 3.2 各区域信令数据情况

对目前获得的数据进行分析，各区域的信令量对比情况如图 6 所示。由于常熟、吴江、张家港、太仓区域的信令总量状况严重偏离基站建设状况（见图 4），因此推断本次所获取数据并不齐全。



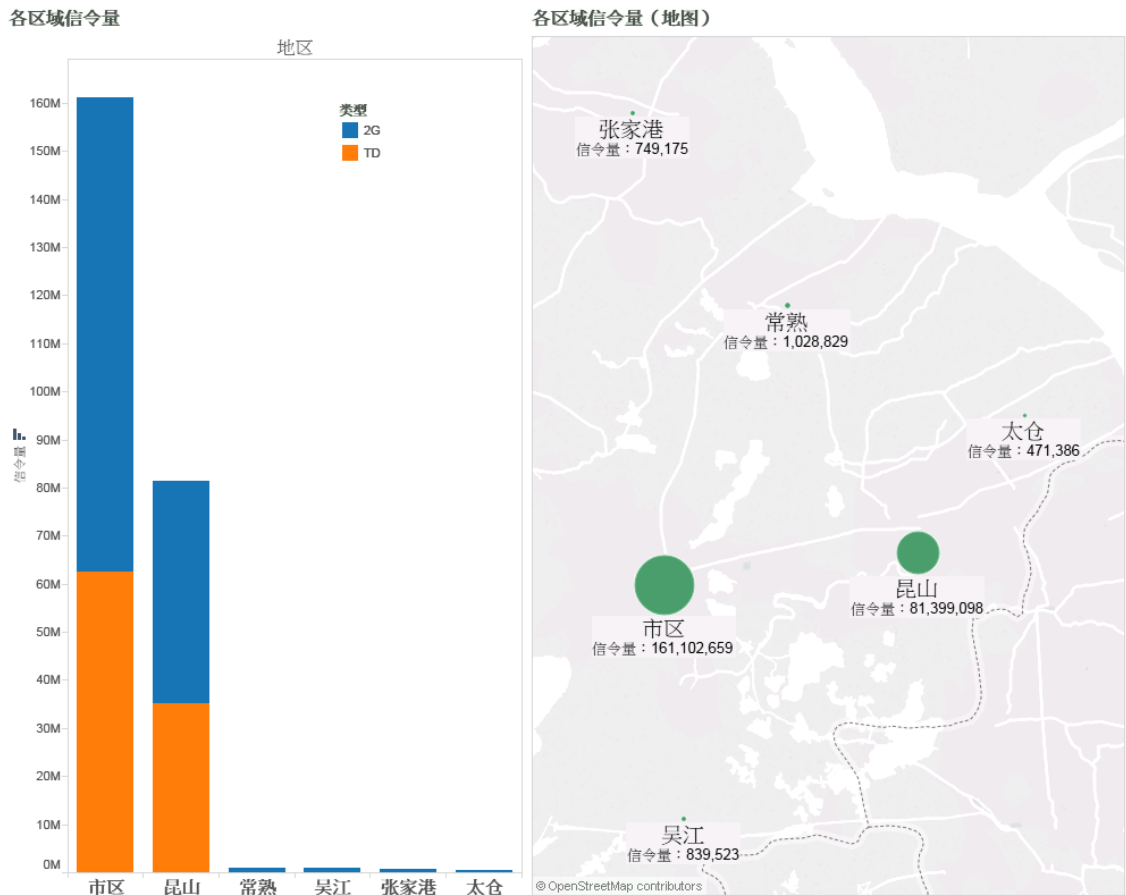


图 6：所获取数据的分布情况

### 3.3 信令数据时变情况

根据手机信令数据的采集量，绘制了信令量时变图，如图 7 所示。图中所反映的数据量变化情况大致稳定，并且符合手机用户行为模式。

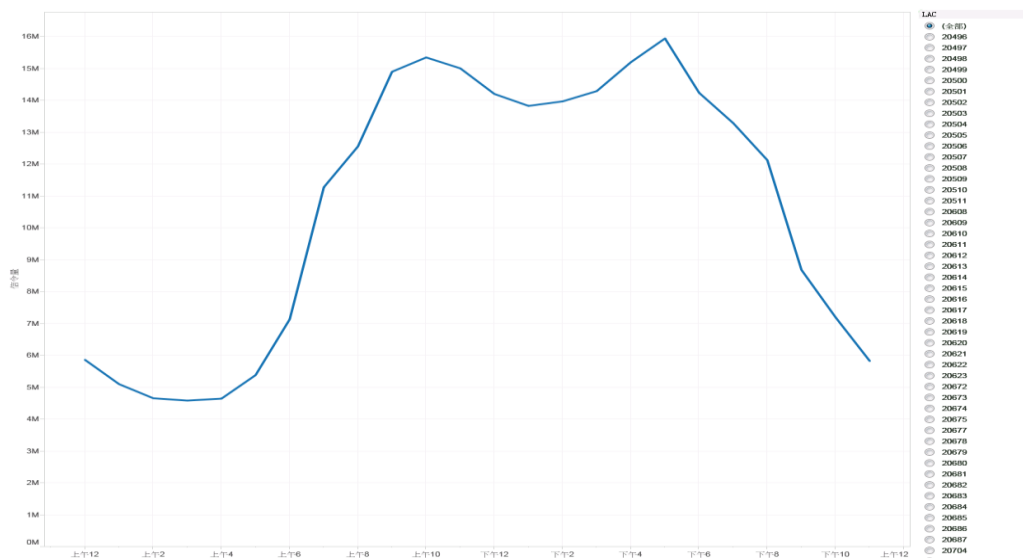


图 7: 信令量总体变化曲线图

# 4 3G 数据的探索

针对目前数据分析，对比各类事件数据的 2G/3G 情况绘制了对比分析图，如图 8 所示。

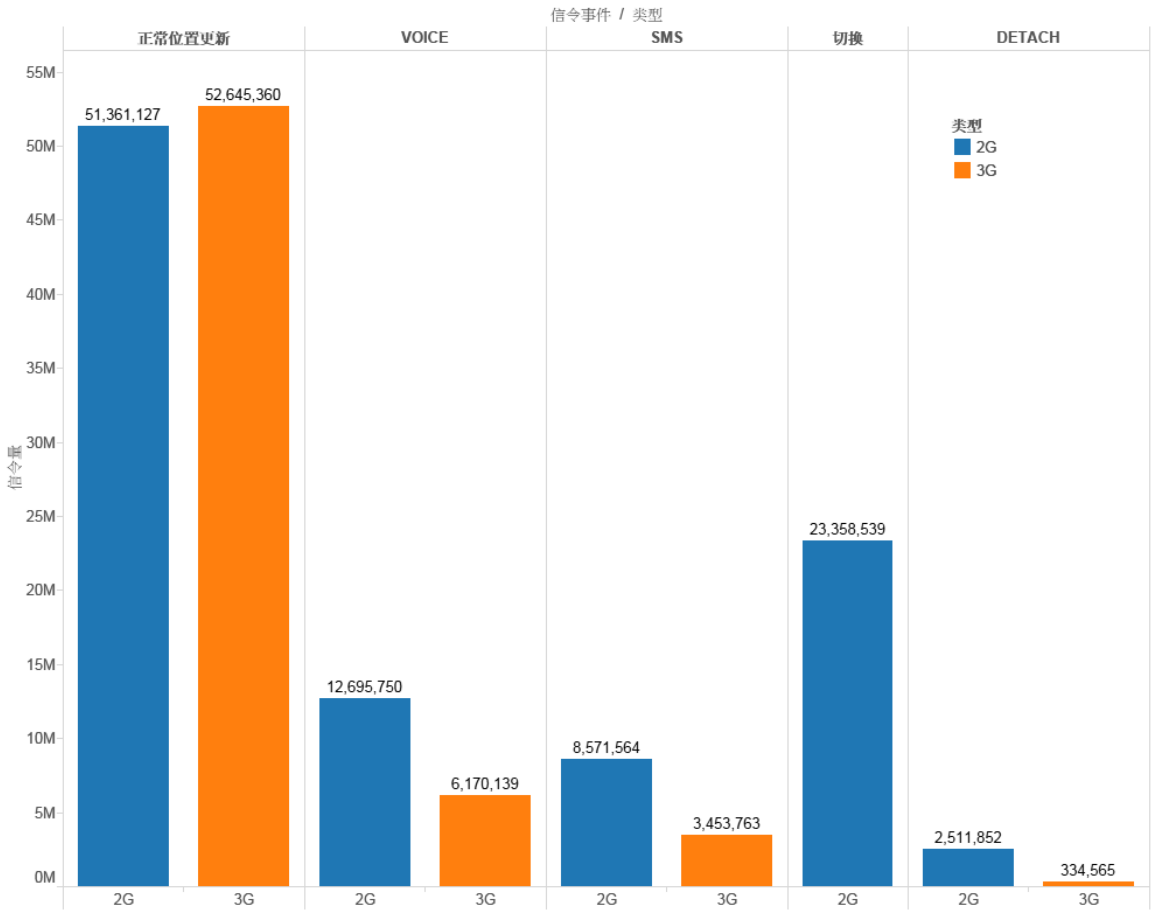


图 8：各类事件的 2G/3G 情况

图 8 所示本次数据中，切换事件不包含 3G 数据。具体原因，需要对数据采集商进行调研。

# 5 需要协调与调研的工作

根据目前所获取的数据文件，还需要协调与调研以下内容：

需要协调获取：

- 1. 完整的信令事件说明；
- 2. 最新的基站信息。

需要调研：

- 1. 本次数据是否为完整的数据，包括：①常熟、太仓等地的数据过少的原因；②夜晚 23 时之后的 1 小时数据缺失的原因；
- 2. 3G 状况：①3G 是否支持通话；②3G 通话时的切换事件信令是否采集。