

答卷

问题一

伪基站常流动于人口密集的区域，以各种名义向一定范围内的手机发送垃圾短信，因此，了解掌握伪基站出行的时空模式，能够帮助执法人员尽早阻止和抓获不法分子，从而更好地维护社会秩序。然而仅仅从垃圾短信中很难确定其对应的伪基站，即无法确定来自同一台伪基站设备的垃圾短信，相同的垃圾短信有可能来自不同的伪基站，同一个伪基站可能不送不同的短信。请从宏观时空分析的角度出发，对垃圾短信数据进行可视分析，揭示伪基站的总体时空活动规律。

1 伪基站的时间活动规律

我们制作的可视化分析系统统计了不同时间段时伪基站所发送的垃圾短信总数，并绘制堆叠图。在这里我们选取其中两天的短信数量变化作为例子进行分析，从图中我们可以看出，垃圾短信从6点左右开始变得活跃，并在8点至10点达到高峰，并在接下来的时间段稍有下降，但一直保持较高的水平，然后在16点至18点又迎来一个高峰。我们注意到这两个高峰时间段都是人们上下班的时刻，许多人在使用交通工具时会有较长的等待时间，违法分子注意并且利用了这个时间段，大量发送垃圾短信吸引这些人群的注意。从18点以后到垃圾短信的数量不断下降，直到22点左右接近谷底。



图1.1 3月10日与3月16日的短信数量变化情况

除此之外，我们设计的系统统计了每日发送的垃圾短信的总数量，并绘制了日历图和热力图，通过圆的大小和颜色可视化地反应该天垃圾短信总数的多少。通过这两个图，我们可以看出，3月4号与3月5号垃圾短信的总数特别多，这恰好是北京召开两会的日期，大量外地人口进入北京，让违法分子有了可乘之机。不

过在此之后垃圾短信数量逐渐减少，到3月15最后到达最低值，这可能是两会以及315打假所带来的监管压力的影响。其次4月初垃圾短信数量相较其它较少，这可能是因为清明期间违法分子也需要回家扫墓或由于上班人数减少，没有了在上下班高峰发垃圾短信的机会，使得违法分子不愿意白白投入大量精力和金钱发布大量垃圾短信，使得垃圾短信的数量减少。作为佐证，我们可以看到周末尤其是周日垃圾短信数量相较工作日减少，这一方面反应了垃圾短信在工作日与休息日的分布特征，另一方面也说明了上下班时间段是违法分子最感兴趣的时间段，应引起相关部门的重视。

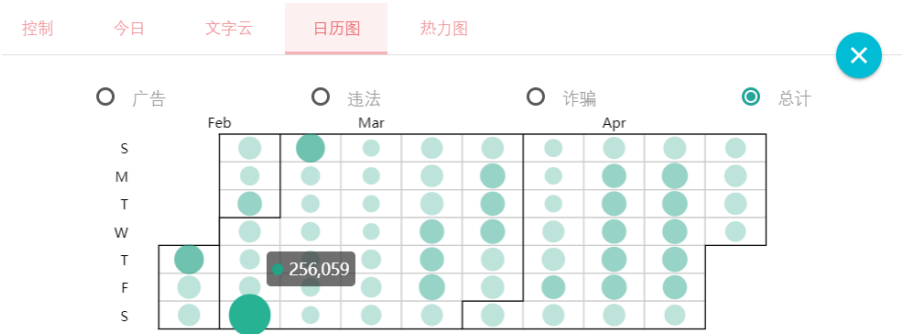


图1.2 垃圾短信数量日历图

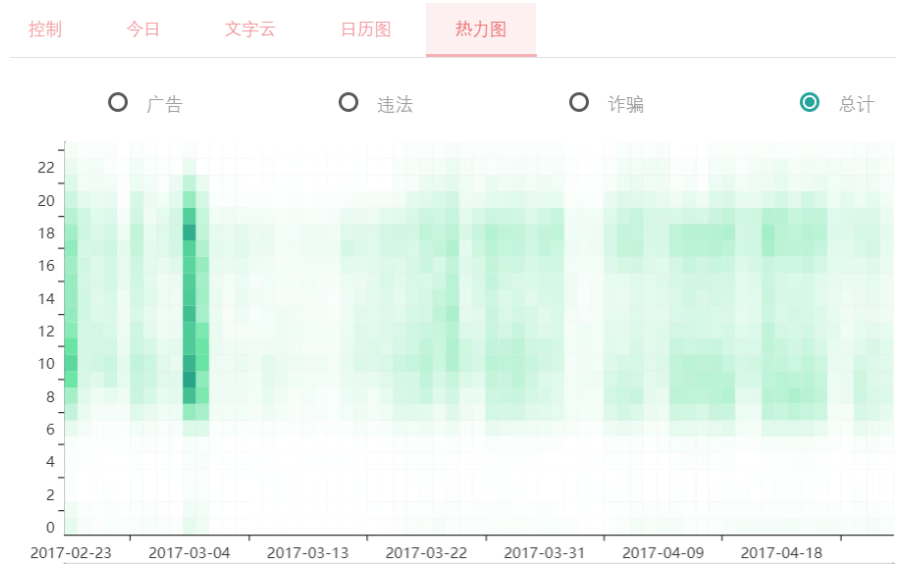


图1.3 垃圾短信数量热力图

2 伪基站的空间活动规律

我们将每天每个时间段的垃圾短信以热力图的形式绘制在了地图上，这样我们便可以通过观察散点的疏密程度了解垃圾短信的空间分布规律。通过图像我们

可以看出海淀区，东城区，丰台区与朝阳区附近短信数量较多。同时我们也可以转成散点图，在散点图中交通要道附近短信相比其他更多，在北京站，北京南站等交通要站附近较为密集，这说明违法分子十分注重在交通便利的地方发送垃圾短信，同时不同时间段垃圾短信的数量会进行一定的变化，说明违法分子有可能在不断改变自己的位置，从而更好地寻找受害对象并逃脱警察的抓捕。

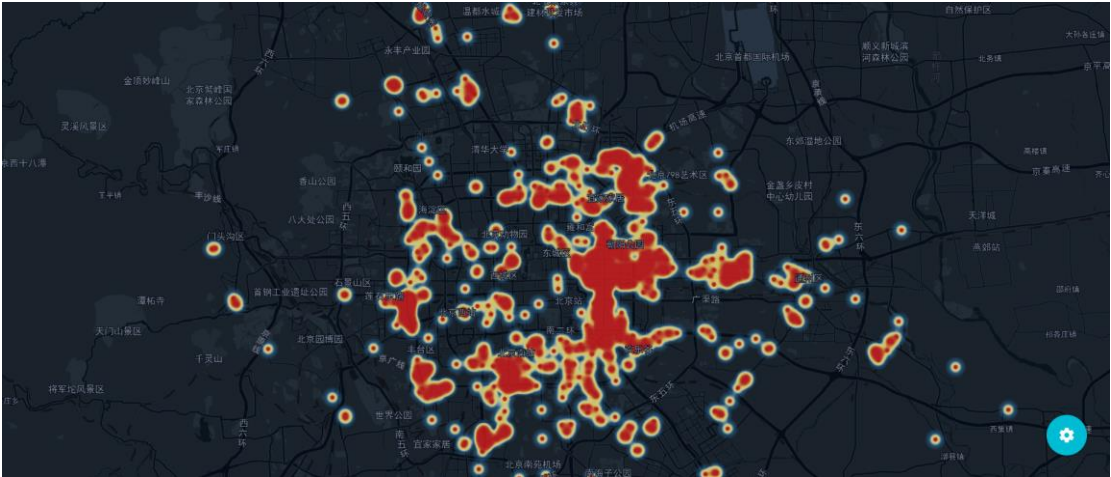


图1.4 某时刻垃圾短信分布热力图

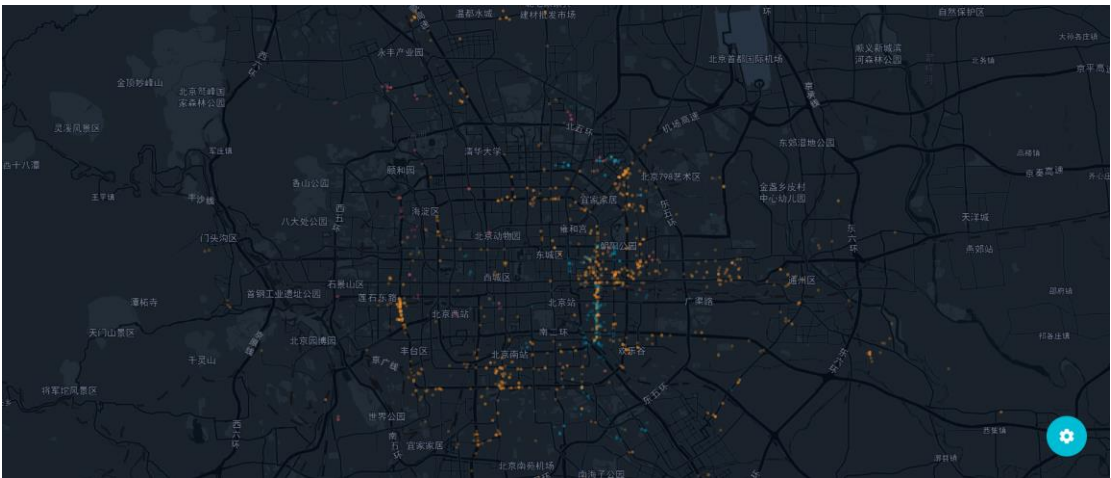


图1.5 某时刻垃圾短信分布散点图

问题二

不法分子通过设置伪基站设备能够发送不同类型的垃圾短信，请尝试对垃圾短信的具体内容进行分类，分类标准不限，例如：按垃圾短信类型可以分为广告、诈骗等等，按垃圾短信对人们的人生经济危害程度可以分为一般、严重等等。请尝试在问题1的基础上进一步分析伪基站发送不同类型垃圾短信的时空分布规律。



图2.3 夜晚垃圾短信高频词云图

系统中的堆叠图中不仅能看到每天不同时间段垃圾短信的数量变化特征，也可以通过不同颜色显示不同短信的数量占比，以及它们单独的变化特征。可以看到，在一天中，违法信息一直占有较大的比重，而广告在主要集中在白天，从6点开始增加，13点到达顶峰，而在21点左右消失。诈骗信息则主要集中在中午10点到12点，以及晚上17点到21点。

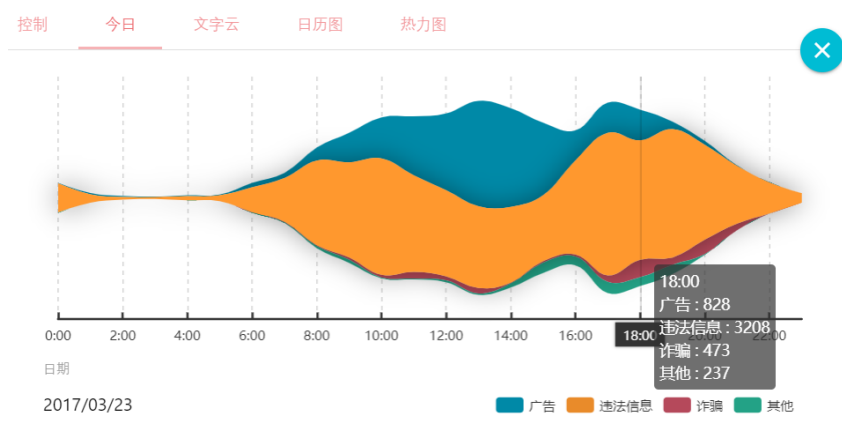


图2.4 不同类型垃圾短信数量一天内的变化情况

系统中的热力图可查看各种类型垃圾短信的分布情况。可以看出，他们在3月5日两会前后十分活跃，在此之后一段时间都有所收敛，但诈骗短信减少的较少。在3月22日左右广告垃圾短信有了一定程度的反弹。而在清明节前后，各种类型的垃圾短信尤其是诈骗短信减少，但此时广告短信减少的程度则较少，由此可以看出不同类型的垃圾短信在不同日期的分布也有区别。



图 2.5 不同类型垃圾短信分布热力图

问题三

伪基站不仅破坏正常电信秩序，危害公共安全，扰乱市场秩序，而且严重损害群众财产权益，侵犯公民个人隐私，社会危害严重。据《人民网》统计，每年通过“伪基站”设备发送诈骗、赌博、推销、中奖等短信近千亿条，伪基站已成为社会一大公害。请结合以上两题中得到的伪基站行为模式，向执法人员提出打击整治伪基站的有效建议和方案，并结合数据分析结果进行说明。

从以上的分析过程中我们可以看出，伪基站的活跃时间主要在6点至22点，高峰时在上下班时间段。在工作日伪基站较多而在休息日有所减少。从空间上的分布上看，我们可以截取一天不同时间段中伪基站分布的散点图来进行分析。可以看到，北京站，北京西站，北京南站这样的交通枢纽附近是伪基站分布的集中区域，并在白天沿交通线分布，在晚上则集中在几个区域。

不同的垃圾短信基站的分布也有所差别。诈骗伪基站主要出现于白天，并主要在东城区，海淀区以及北京西站，北京站附近。广告伪基站分布的区域则更为狭窄，主要是在东城区附近。违法信息伪基站在白天分布较广，而在晚上则主要分布在东城区，朝阳公园附近，这可能是大量发送违法信息垃圾短信的人员的居住地。因此，我们建议有关部门白天加强在北京站等交通枢纽以及各主要交通要道的监察，对疑似装载伪基站的交通工具进行抽查。晚上则对东城区的垃圾短信进行重点监测，对可疑人员进行追踪和询问。

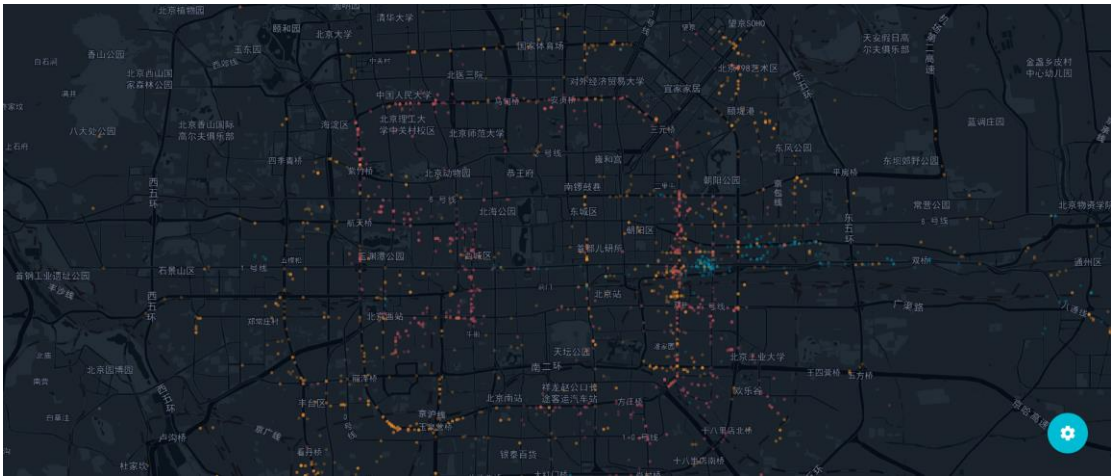


图3.1 白天不同类型垃圾短信分布

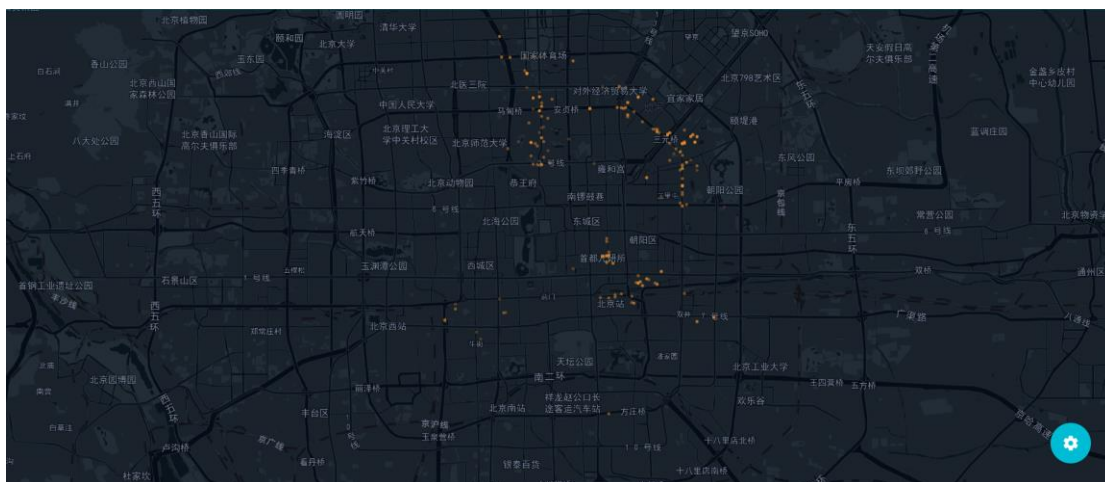


图3.2 夜晚不同类型垃圾短信分布