1. 使用的是 MATLAB
2. 关于应用的数据集是模拟生成的 按照设定的概率 微服务的候选对象数量等
3. 用户和SBS 是真实数据集 **地理位置** 目前已经找到

这个实验求的是 整个在应用上的处理时间最少 包括 in+exe+out

所以没有用户对于应用请求频率这个概念 而且研究的是 只有一个应用 实际数值设定是 5个微服务组成的一个应用，每个微服务有[2,8]个候选人

现在的问题是：

还需要应用请求频率

另外一篇使用的数据集：

三大云服务提供商 VM type&price

User center

网络延迟 真实数据集

仿真设置：

1. 8种工作流结构
2. 各个服务的处理时间
3. 使用Facebook subscribers的数据模拟应用请求的规模 是facebook的1/1000，模拟拥有百万用户的应用程序提供商，是从82个用户中心随机选出30个，每个用户每天发出25个请求， Accordingly, the application request rate spans from 52 to 304 requests per second

OK：**经纬度** 相差0.01就是1000m了

边缘服务器位置

用户位置

应用数 自己设置 5

每个应用设计的服务 自己设置 [3,8] 只是线性最基本的就好

总共服务数 自己控制 结合上面两个

用户对于应用的请求 要假设**百万级用量**吗？ 一个用户一天对于一个应用请求频率 假设在5就好了把

对于应用要设置一下 请求频率高的请求频率低的，是从所有的用户当中选取百分之多少的用户来发起请求。

**最好具体化一点：**

这样子比较好想像啊 还是从最初举的例子下手，支付和拍照软件

那**基本实验数据获取**完成了 主要是代码了 想好使用什么语言，虽然我还是喜欢C++，但要用Python的包

实际应用： 6个应用每个应用5个模块 复用之后 一共20个模块

请求频率设定： 5 10 15

大事情：出现一个逻辑上的问题就是：一天这个人都不会待在商业区的，按道理应该是有变化的，但是强行解释的话，人是流动，在每天这个时间段里总有人是在商业区的，所以按道理以后考虑时间会更合理。

聊天（都要 高等） 1 13 14 15 7

游戏（生活 中等） 2 14 16 17 8

音乐（都要 但是 低 分布多） 3 14 16 20 9

拍照（旅游 但请求高）4 13 14 17 10

支付（商业 请求最高）5 15 13 18 11

媒体（都要 中等） 6 14 16 19 12

加了离散度权值之后就是 1和2的优先级发生了变化，但事实上应该是分布更广泛的 要优先放 原因是超过平均值的才选择了 那些都是相同的 所以没有平均值，改成 大于等于平均值的 要算离散度，聚类的依旧是大于平均值

各自开始与结束

13是面部识别3

14是图像渲染5

15是信息认证2

16是音乐3

17是VR或者其他2

18是支付接口1

可以用一段话 解释不同应用的请求频率等等然后再举例子说明。

服务模块 有出现的概率 随机生成 从属于不同的应用 就按照上面的来

40个ES(随机挑选 还是要按区域的啊) 500个USER

一个用户一天对于一个应用请求频率

对于应用用户的分布情况也要模拟的？

数据处理：

1. 随机选 就选一次哈 选完之后 可视化一下
2. 根据可视化的结果 对于覆盖用户范围有概率选择+用户请求频率按概率在某一个范围内出现

**AP优点**有：

1. 不需要制定最终聚类族的个数

2. 已有的数据点作为最终的聚类中心，而不是新生成一个族中心。

3. 模型对数据的初始值不敏感。

4. 对初始相似度矩阵数据的对称性没有要求。

5. 相比与k-centers聚类方法，其结果的平方差误差较小。

做法记录：

获得每个用户对于每个模块的请求频率之后，直接用AP来算，P是一个固定的值，结果就是怎么都无法收敛，后来一想肯定的，没有把请求频率这个权值考虑进去，仅仅是距离进行聚类，那岂不就是讲用户聚类一个道理，分散太广泛了不行的

现在讲p变成weight后，每一个点就是自己一个类了变成

错误在于：simi中中间值也就只有-0.0073105 结果我的weight都是十位数的，那肯定自己归自己啊。。

变成0.0几后这样子之后 还是出现原来的问题 没有办法收敛

应该还是因为太多了

~~所以决定 超过一般请求以上的点 才聚类 否则没有必要聚类~~

变成0.几后，实现了聚类，但是只有两个点，

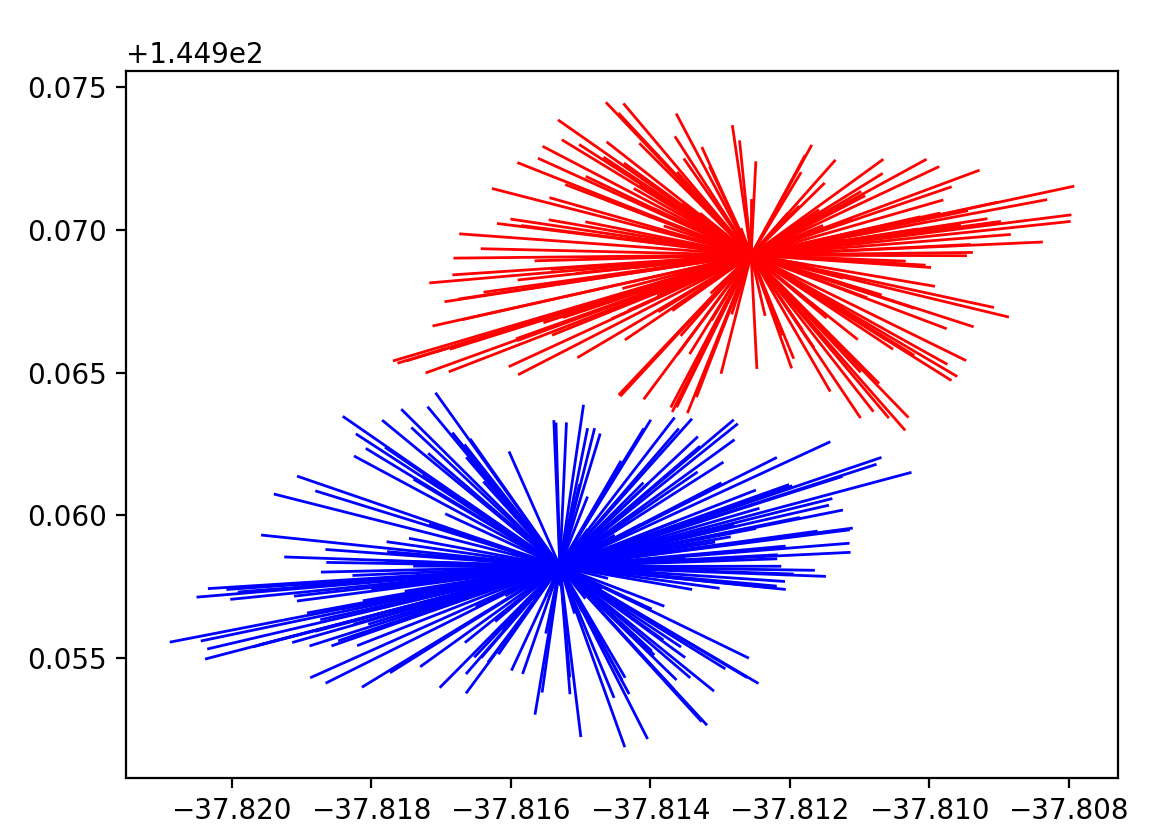


图 1 17的聚类结果

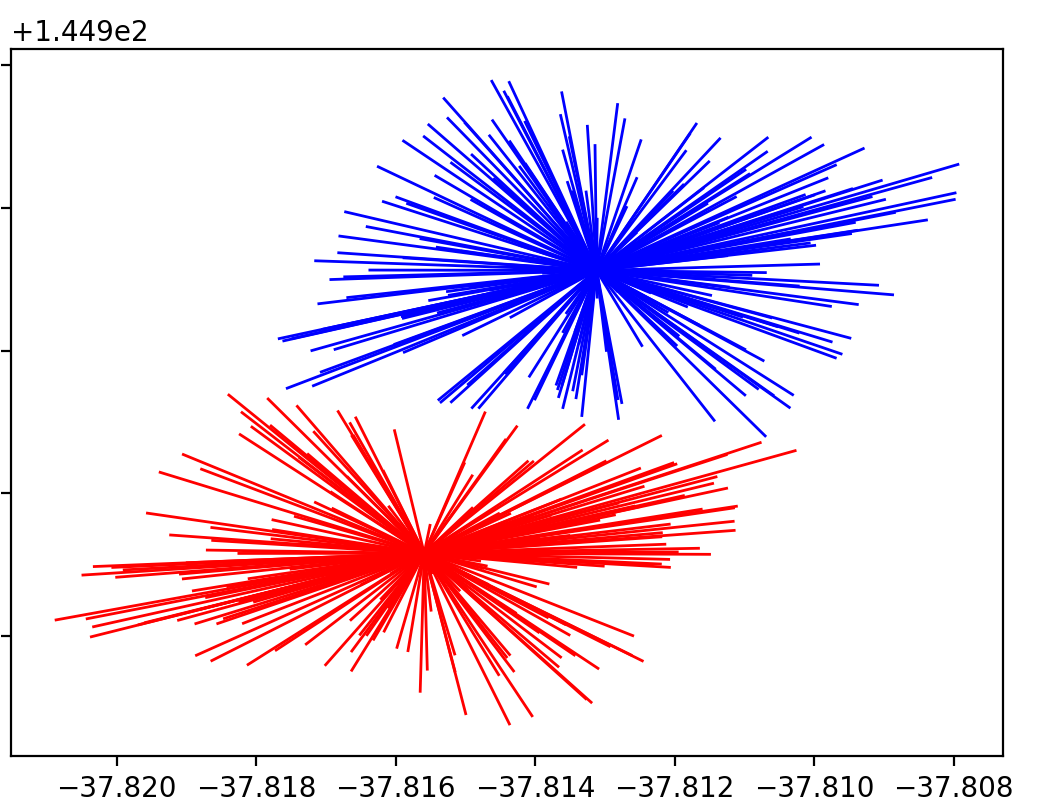


图 2 18的聚类结果

但问题是18是支付接口，但道理在商业区应该有一个 其他地方无所谓的

再大 就又收敛不了了

参数再调小 就只能收敛城一个了

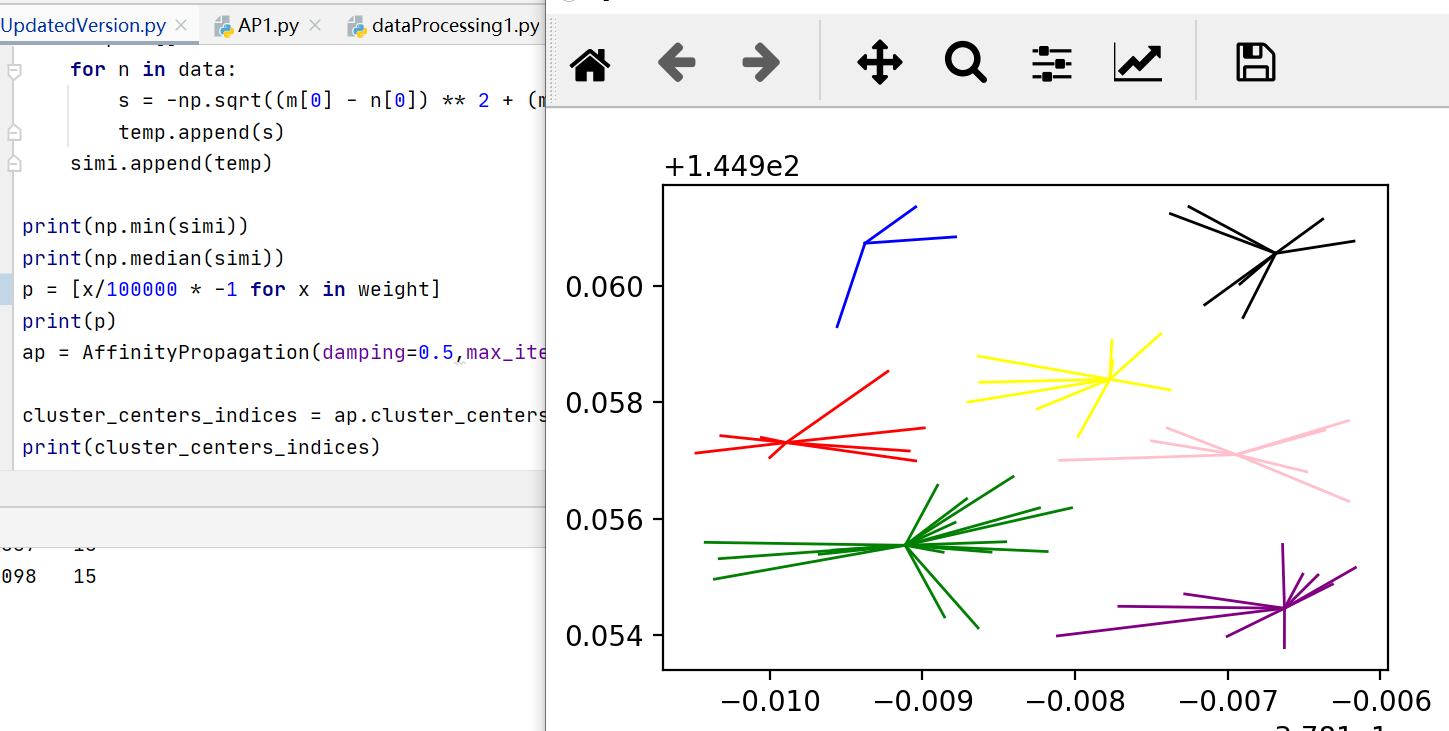


图 3 经过筛选 10000

经过筛选 大于均值的是有68个 将p调到了除以100000之后，举出来了7个类，但问题就是18支付接口应该只有商业啊？？问题出现在：这已经是在商业区里面进行聚类了- - 如果按在大区上面聚类确实是只剩下了一个没错

尝试了s13模块 但道理会在商业和旅游两个地方，但/1000又收敛失败

/10000后

