

数据库大作业

计 24 2012011335 柯均洁

一、 实验内容

● 输入：

■ POI点的集合

➤ 经纬度(Longitude, Latitude)

➤ 名字name

■ 矩形查询区域

■ K

● 输出：

■ K个出现频率最高的tag

K nearest neighbors to the query location that contain all keywords

二、 运行方式

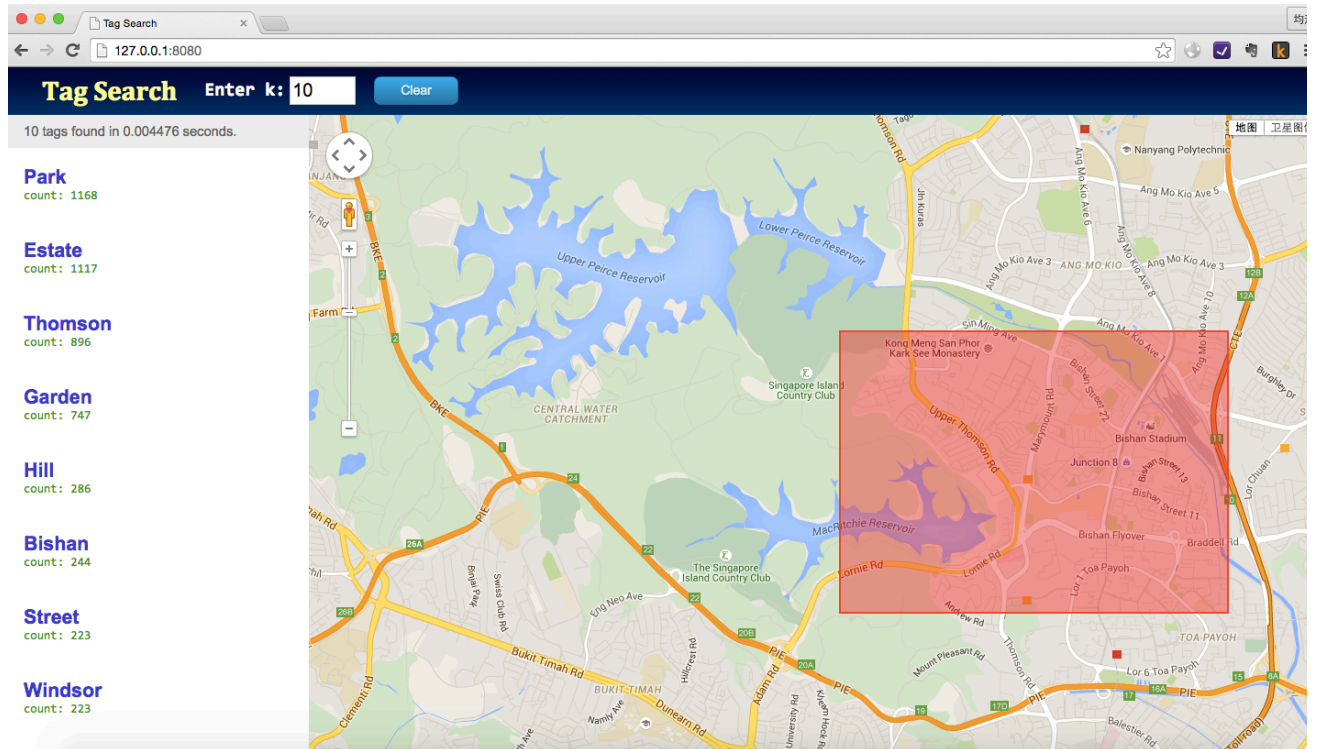
终端运行：

1. `cd map_db/`
2. `make clean py`
3. `mv db.so ../server/server`
4. `cd ../server/`
5. `python manage.py runserver`

在浏览器访问：localhost:8000

三、 使用说明

点击地图设定查询区域，再点击一次确定查询区域。输入 k。就实时显示查询区域内出现频率最高的 k 个 tag。



四、 项目实施

1. 用 c++实现查询函数，并用 Python.h 进行封装
 - ./map_db/db/search.cpp
2. 用 django 实现前端
 - ./server/
3. 用 js 来调用 google api
 - ./server/static/js/main.js

五、 实现细节

- 数据库建立
 - a) 读入 json 文件中的地点，并调用 RTree 的 Insert 建立一棵 RTree
 - b) 对地点名称 name 进行分词，建立单词词典，并进行编号，每个地点保存一个单词序号的集合
- 查询

根据用户选择的矩形访问 RTree，取出命中的节点，统计单词出现的频率，并用堆取出前 k 个出现频率最高的单词
- 结果显示

将出现频率最高的 k 个单词，及其频数返回到前端，并进行显示