doi:10. 3969/j.issn. 1672-4623.2023.02.017

顾及约束的交互式热力图服务研究与应用

吴瑞龙[1]，王明省[2]，何华贵[2]，张鹏程[2]，林　鸿[2]，龚　磊[2]
（1. 武汉市自然资源和规划信息中心，湖北 武汉 430014；2. 广州市城市规划勘测设计研究院，广东 广州 510060）

摘　要：针对现有热力图可视化方法无法满足基于约束条件实时交互生成问题，提出一种顾及约束条件的交互式热力图服务实现方法：首先解析服务请求的约束条件，包括空间约束和非空间约束，服务端对空间点要素数据进行初选，接着对空间位置较近或重叠的点要素按距离容差去重并累加权重，然后利用DBSCAN密度聚类算法进行空间聚类，计算相同类簇样本集的几何重心位置和权重，最后将合并后的结果集返回前端进行渲染生成热力图，依据此方法设计了交互式热力图服务接口，基于广州市商事登记主体数据分布应用结果表明：此方法可满足约束条件下交互式热力图服务的需求，保证点要素在空间分布密集程度准确性的同时高效地生成热力图，具有一定的适用性和应用价值。

关键词：约束条件；交互式；热力图服务；DBSCAN；空间聚类
中图分类号：P208　　**文献标志码：**B　　**文章编号：**1672-4623（2023）02-0085-04

Research and Application of Interactive Heatmap Service Considering Constraints

WU Ruilong[1], WANG Mingsheng[2], HE Huagui[2], ZHANG Pengcheng[2], LIN Hong[2], GONG Lei[2]
(1. Wuhan Natural Resources and Planning Information Center, Wuhan 430014, China;
2. Guangzhou Urban Planning & Design Survey Research Institute, Guangzhou 510060, China)

Abstract: In this paper, we proposed an interactive heatmap service implementation method considering constraints, aiming at the problem that the current heatmap visualization method couldn't meet the real-time interactive generation based on constraints. Firstly, we analyzed the service request constraints, including spatial constraints and non-spatial constraints, and used the server to select the spatial point element data preliminarily. And then, we deleted repetition for the point elements with close or overlapping spatial positions according to the distance tolerance, and accumulated the weight. After that, we performed spatial clustering based on DBSCAN algorithm, and calculated the geometric center and weight of the same cluster sample set. Finally, we returned the combined result set to the front end for rendering to generate heatmap. Application for data distribution of Guangzhou business registration enterprises shows that this method is applicable and valuable, which can meet the needs of interactive heatmap services under constrained conditions and ensure the accuracy of point elements in the spatial distribution density meanwhile efficiently generate heatmaps.

Key words: constraint condition, interactive, heatmap service, DBSCAN, spatial clustering

目前关于热力图研究多侧重于热力图的前端渲染和优化实现，依赖于前端的交互信息，如前端的屏幕显示范围或者前端的地图缩放级别，这些研究都是基于静态专题图，利用现有GIS服务器制作专题热力图，然后发布为地图服务，服务端的处理一般只是依赖于属性的查询过滤，将查询后的数据一次性返回渲染，没有做聚合的处理来压缩数据，热力图的渲染效率还比较低下[1-15]；另外，现有研究更多是基于无约束条件的热力图，针对时空大数据，较少考虑时空约束条件下交互式的实时生成热力图的研究，不能满足用户特定需求[16]，难以通过服务的形式提供交互式的实时生成热力图，比如针对用户感兴趣的某个特定时间段，特定时间区域，不能实时的生成热力图。

为此，本文提供一种顾及约束的交互式热力图生成方法，该方法允许在给定约束条件下对空间点数据进行过滤，实现数据的动态筛选，并利用距离容差和DBSCAN聚类算法按照其空间分布进行聚合，达到快速生成热力图的目的，同时设计了交互式热力图服务的RESTful接口，能够有效满足交互式热力图服务的需求，具有一定的应用价值。

1　总体技术方案

本文技术方案采用B/S模式架构，统一了客户端热力图服务访问的请求。服务端根据请求进行实时计算并返回热力图计算结果，实现热力图可视化和计算

收稿日期：2020-01-09。
项目来源：智慧广州时空信息云平台建设项目（GZIT2016 – A5 – 147）。

的前后端分离。

1）由客户端发起带约束条件的热力图服务请求，服务端对约束条件进行解析，包括空间约束和非空间约束，服务端按照约束条件对空间点要素进行初步筛选过滤。

2）对空间位置较近或重叠的点要素按给定的距离容差去重并累加权重，再利用DBSCAN密度聚类算法进行空间聚类，计算相同类簇样本集的几何重心位置和权重。

3）将结果合并转换为GeoJSON格式数据，返回给前端框架如OpenLayers、Leaflet进行加载渲染实时生成热力图。总体技术方案流程如图1所示。

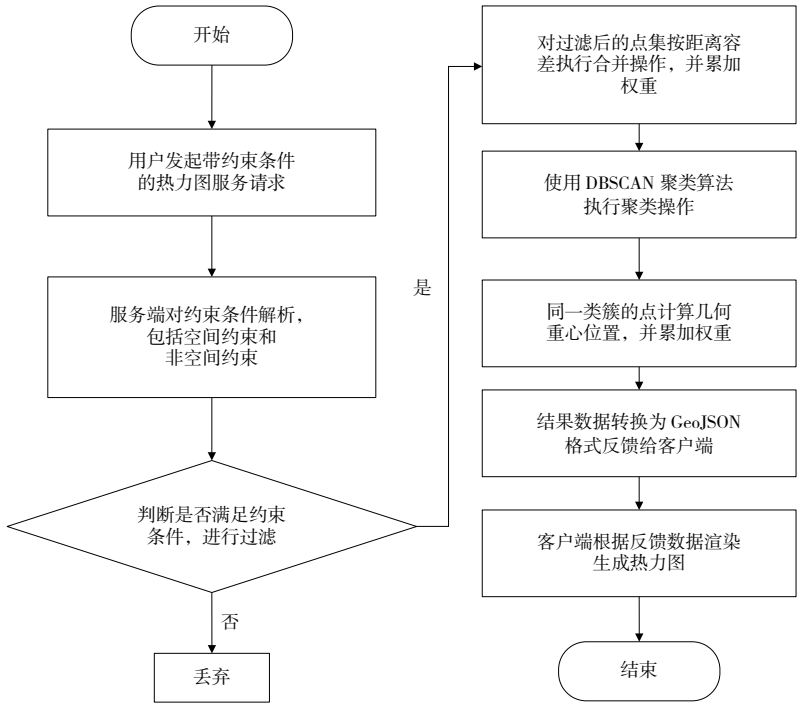


图1 总体技术方案流程图

2 关键技术

2.1 按照距离容差合并

根据地理学第一定律，空间地物之间的相关性与地理位置距离有关，地理位置较近的相关性越大，反之相关性越小，因此对于地理空间样本点距离较近或重叠的情况，本文首先对样本数据按照给定的距离容差进行合并，累加其权重，在服务端通过距离容差对空间位置较近或重叠的数据进行排除，实现了数据的一次性过滤，排除部分数据干扰，减少了数据计算量，有效提高计算效率。具体计算方法如表1所示。

2.2 DBSCAN空间聚类

DBSCAN是一种基于密度的空间聚类算法，可以按照空间距离对点样本进行聚类，该算法能够发现任意形状点簇，聚类速度快，同时也表达了空间点的聚集程度，适合用于本文的实时热力图可视化表达，该算法主要涉及2个参数：一是领域半径Eps，该参数作为类簇搜索的距离依据，通过初始中心点按照邻域半径进行判断，形成一个类簇的闭包；另一个是MinPts，该值作为聚类的密度约束，是在给定点邻域内的最少点个数。结合总体技术路线，本文后端服务基于DBSCAN聚类的处理流程如图2所示。

表1 按照距离容差合并算法描述

说明	通过距离容差对空间点集进行合并
输入	初始点集 D_1，容差参数 ε_d
输出	合并后的点集 D_2
1	MergePoints(D_1, ε_d)
2	For p_i, p_j in D_1
3	计算两点距离 d_{ij}
4	if $d_{ij} <= \varepsilon_d$
5	取点 p_i 加入集合 D_2
6	累加权重 $w_i = w_i + w_j$
7	End if
8	End for
9	Return D_2

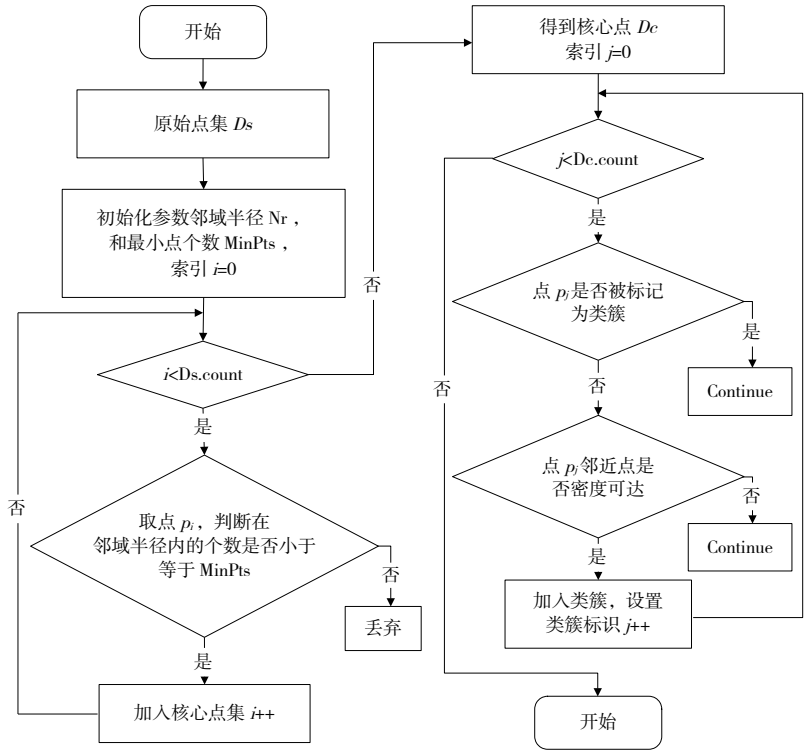


图2　服务端空间聚类处理流程图

2.3　相同类簇计算地理重心及权重

为了更准确地表达聚类后样本的密度分布和聚集程度，需要计算同一类簇的几何中心坐标并对同类簇的点集权重累加，计算方法：遍历相同类簇点集中的每一个类簇点集，设类簇为 k 的点集为 $P^k=\{p_1^k,p_2^k,p_3^k,\cdots,p_n^k\}$，其中任意一点为 $p_i^k=(x_i^k,y_i^k,w_i^k)$，$x_i^k$，$y_i^k$ 为点的坐标，w_i^k 为权重字段值，则计算该类簇的几何重心坐标以及权重值为 $\left(\overline{x_k},\overline{y_k},\overline{w_k}\right)$，计算公式如下：

$$\begin{cases}\overline{x_k}=\sum_{i=1}^{n}w_i^k x_i^k/\sum_{i=1}^{n}w_i^k\\ \overline{y_k}=\sum_{i=1}^{n}w_i^k y_i^k/\sum_{i=1}^{n}w_i^k\\ \overline{w_k}=\sum_{i}^{n}w_i^k/n\end{cases}\tag{1}$$

2.4　交互式热力图服务接口设计

基于以上技术路线，本文参照标准WMS规范设计了交互式热力图服务HMS规范，服务返回结果集使用GeoJSON格式，满足地理信息服务数据交换与共享的需要，可适用于主流的WebGIS前端渲染框架，详细定义如表2、3所示。

热力图服务请求示例如下：

http://localhost/mapserver/HMS?VERSION=1.0&REQUEST=GetHeatMap&LAYER=’SSDJ’&PROPERTYNAME=name&SQL_FILTER=” name=´Blaine´” &SPATIAL_FILTER=‘Polygon((3 1,4 4,2 4,3 1))’

表2　热力图服务HMS规范列表

规范项	规范说明
协议	HTTP
请求方式	GET
状态码	遵循HTTP协议状态码
返回格式	GeoJSON
实现操作	GetHeatmap

表3　GetHeatmap操作请求方法实现参数列表

参数名	参数描述	是否必须
SERVICE	服务名称，即HMS	是
REQUEST	GetHeatmap	是
VERSION	1.0	是
DISTANCE	容差容差	是
SPATIAL_FILTER	空间过滤条件，WKT形式参数	否
SQL_FILTER	基于SQL形式的非空间过滤条件	否
LAYER	所请求生成热力图的图层名	是

3　应用案例

商事登记主体数量一定程度体现了区域经济发展水平的高低，通过对商事登记主体的大数据挖掘分析，有助于了解地区经济的活跃情况和产业发展状

况，如何直观动态地对产业分布现状进行可视化表达有着重要意义。本文实验数据来源于广州市政府数据统一开放平台，共约10万条，数据格式为CSV，通过脱敏清洗处理后包含的字段有：注册号、名称、主营项目类别、经营范围、成立日期等，使用高德地图地理编码Web服务API，利用名称字段将其转换为地理坐标落图，数据源部分示例如表4所示，全部数据落图效果如图3所示。

表4 商事登记主体数据部分样例

序号	注册号	名称	主营项目类别	经营范围	成立日期
1	440126000423138	广东晾的智能科技有限公司	001M01073	电子、通信等	2014/03/14
2	440126000426851	广州浓兴贸易有限公司	001F01051	商品批发贸易等	2014/03/26
3	440126000435944	广东龙大实业投资有限公司	001L01072	企业管理服务等	2014/04/22
4	440126000443792	广州哲丰科技有限公司	001M01075	图书处理技术开发等	2014/05/14
5	440126000475144	广州永毅经贸有限公司	001F01052	化工产品批发等	2014/06/11

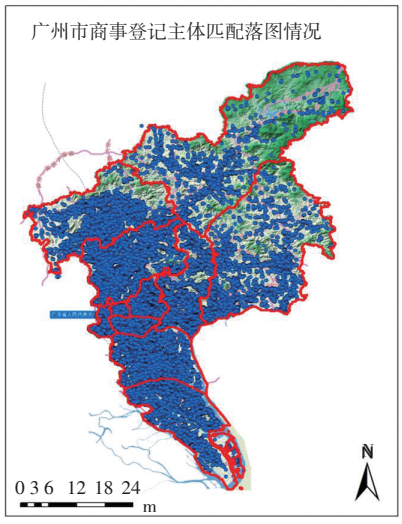


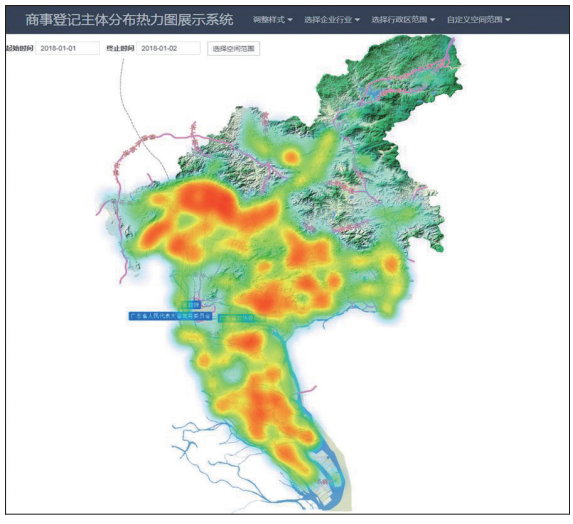
图3 商事登记主体落图示意图

本文应用系统开发基于Windows10环境，前端使用OpenLayers3框架，后台开发及算法实现基于Java JDK1.8版本，热力图服务HMS实现基于Vert.x的Http模块，空间数据的存储和运算使用开源数据库PostgreSQL和PostGIS插件。系统以B/S架构模式开发，支持主流浏览器访问，系统部分截图如图4所示。主要实现的功能有：

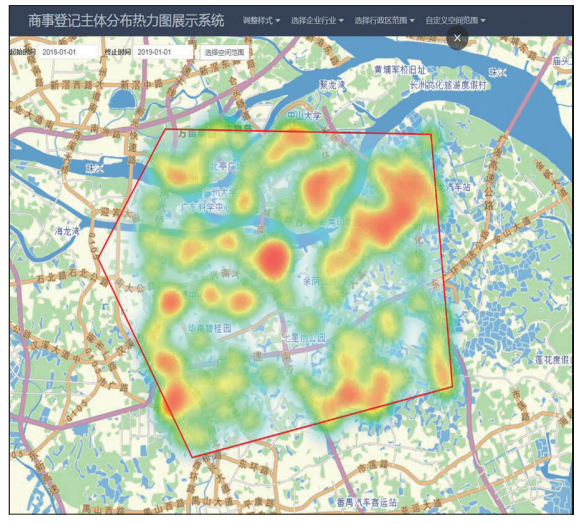
1）按照给定空间范围实时显示热力图，包括任意拉框范围、矩形范围、圆形范围。

2）按照给定属性条件实时显示热力图，包括商事登记主体的主营项目类别，经营范围类型，成立日期等。

3）热力图样式设置，包括热力图半径大小调整和颜色设置。



a 选择某个行业实时生成全市范围热力图



b 绘制任意多边形实时生成该范围内热力图

图4 部分系统截图

4 结 语

本文从实际应用场景出发，在顾及约束条件下服务端对空间点数据进行动态筛选并根据距离容差进行合并，实现一次过滤；基于DBSCAN的空间聚类算法，按照样本数据的空间密度聚类，实现二次过滤，同时计算同一类簇的重心坐标和权重值，既保留了样本数据的聚集程度信息，表达了其空间分布情况，又减少了数据的计算量，提高了计算效率，达到实时动态

（下转第102页）