

VIS2020学术会议的推特数据可视化

数据描述与分析

在本任务中，我们需要对会议相关的推特进行分析，从中获得关于会议的基本信息。

数据分为两部分，一部分是推特数据，一部分是用户数据。

推特数据的属性有：创建时间、发推用户名、推特内容、点赞数、回复数、转发数、相关标签。我们对点赞数、回复数和转发数这三个属性做了加权统一，得到一个新的属性，为推特的热度。其中：

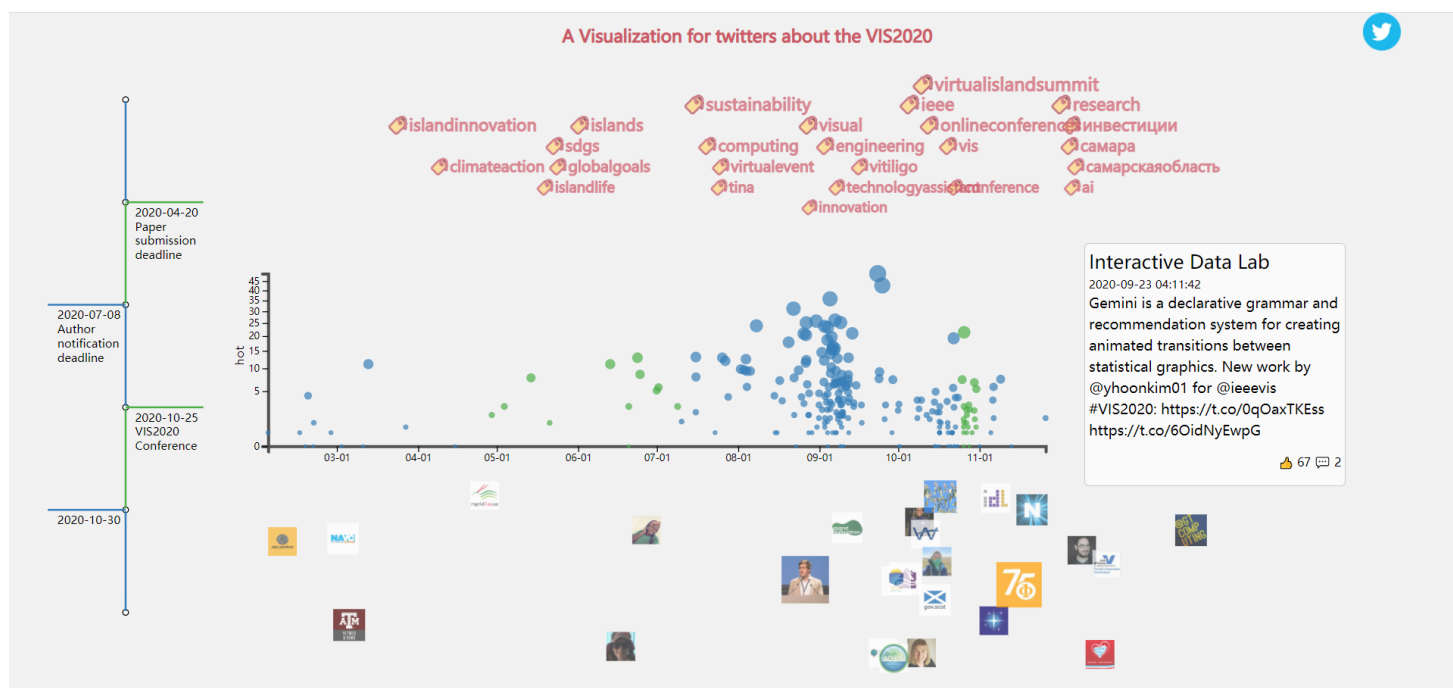
- 创建时间、热度为定序型变量
- 发推用户名、相关标签为定类型变量

用户数据的属性有：用户名、粉丝数量、用户头像。其中：

- 用户名、用户头像为定类型变量
- 粉丝数量为定量型变量

在进行数据处理后，我们得到了用户发推时间平均，以及不同tag的热度，从而获得系统设计的各个参数。

项目设计与展示



主体描述

系统分为四部分，以时间为主轴展开。

第一部分为tag的散点图，展示了在这个会议期间热度较高的话题，话题的热度通过字符大小展示，讨论越多的话题字号越大，而话题的横坐标是由话题被引用的平均时间决定的，在时间轴上横向分布形成对这些推特内容的“词云”。

第二部分为推特的散点图，每一个点代表一条推特，点的横坐标对应着发推时间，而点的纵坐标代表这条推特的热度，为了使得所有的点在纵向上成均匀分布，我们设计了不均匀的纵轴刻度。在该图内部的交互中，当把鼠标悬停在某一条推特上时，界面右侧会显示该推特的内容，包括发推用户、推特时间及内容、点赞量和回复量，我们在该部分极大程度地还原了twitter.com上的视图。在时间轴上进行滚轮滚动会放缩范围，从而为用户按需展示不同时间段的推特散点图。

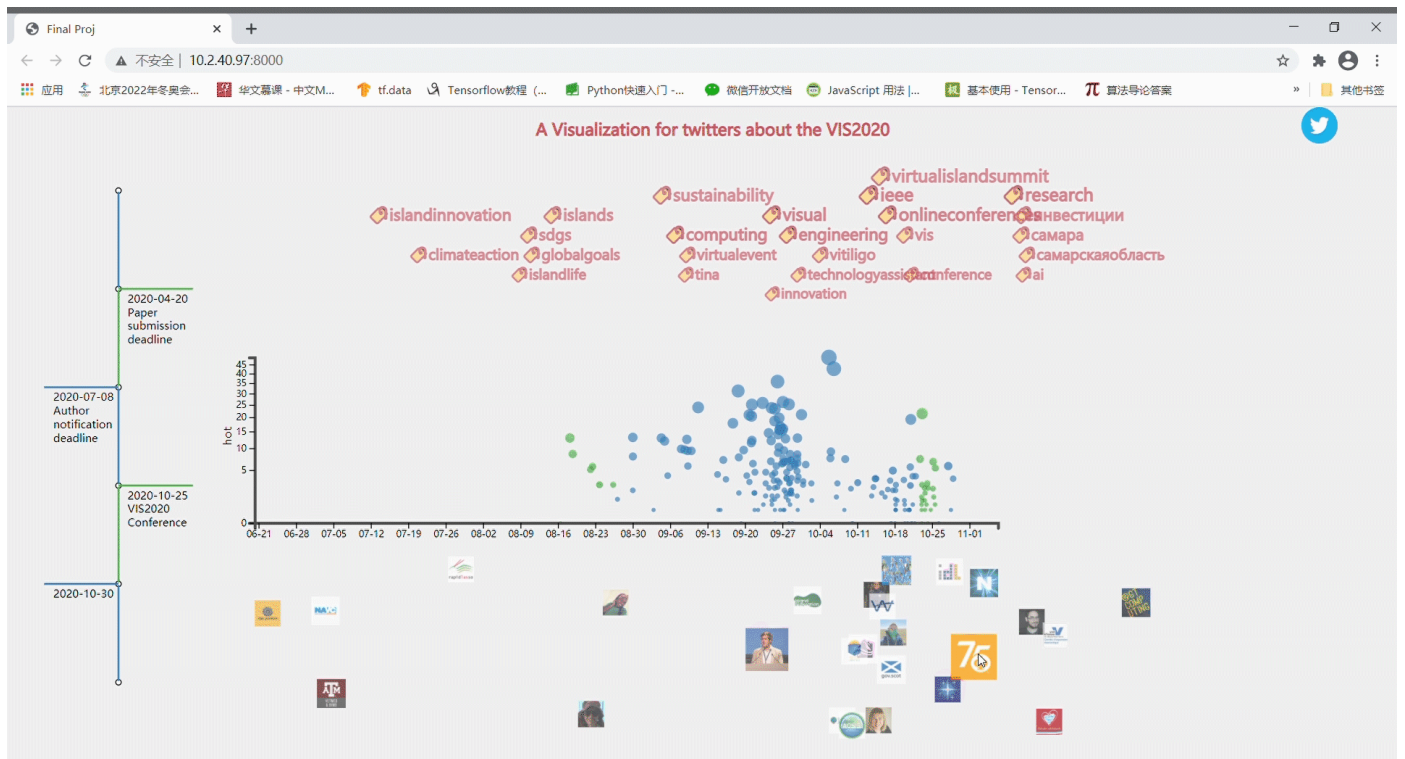
第三部分展示了在该会议中活跃程度较高的用户，图片是该用户的头像，图片大小表示了该用户的影响力大小，这些图片的分布也类似于tag的散点图，横坐标由发送的相关推特的时间决定，从而在时间轴上横向分布。鼠标悬停在某个用户头像上时，会在头像上方展示该用户的名字。

第四部分为时间轴，展示了可视化会议的时间线，包括提交论文的deadline等时间点，并且不同时间段用不同颜色区分，对应时间段的推特也用了相应颜色在第二部分的散点图中标注出来。

交互

三部分的数据之间联系紧密，一条推特可能对应多个标签，一个标签也可能对应多条推特，某个用户也有ta关心的不同话题，基于以上特征，我们加入了以下的交互：

1. 点击第一部分的某个标签，在图二会高亮带有该标签的所有推特，图三会高亮关注这个标签的所有用户，如下图



可视化结果分析

1. 主要推特集中在会议进行时间的前后，且这些推特的关注度也较高。由于2020年的会议均在线上进行，所以很多推特中都带有了#online这个关键词。
2. 热度最高的推特是Interactive Data Lab所发的两条有关可视化工作的成果，分别是关于Gemini和Boba可视分析方法，但是该用户在整个会议期间活跃度并不高，总共只发了三条推特。
3. 发推最多的用户是James Ellsmoor，他是可持续发展和可再生能源中的佼佼者，是太阳能国家首脑的联合创始人和董事。
4. 热度最高的话题是#virtualislandsummit，即虚拟岛屿峰会，和#sustainability（可再生能源的持续性），以及#ieee。
5. 用户也有自己倾向的讨论话题，例如IEEE ComputerSociety用户发的推特中包含的#ieee标签较多，可以看出该用户倾向于此话题。
6. 通过和会议进程的联系，我们可以看到，在论文递交的deadline附近，推特上有很多网友发表了自己撰写论文时的想法和遇到的困难，也有网友发表了自己撰写的论文的主题，例如可视化在数据科学中的应用、互动社交网络集群等。在会议进行中，有些权威机构的官方推特发表了会议中的结论，例如新的方法的概述和应用，也有一些与会专家在推特中发出自己的演讲的通告，同时也有一些大公司迅速将研究成果应用在自己的产品中，也有团队在会议中获奖后发出的庆祝推文。

关于本工作的进一步讨论

1. 本工作最大的问题是数据集的收集。我们发现，#VIS2020可以代表两个会议，即可视化会议和虚拟岛屿峰会。我们开始时并未意识到#VIS2020会对应两个完全不同的会议，直到最后分析结果的时候

才发现，这个bug直接导致我们无法看到希望的有关可视化会议的内容。如果更换一个会议，可能展示的效果更好。

2. 我们的时间轴与其他视图没有产生交互，这一点在后期可以进行补充。
3. 在展示用户的时候，我们依照用户的粉丝数进行了筛选，过滤掉了一些粉丝数极少的用户，从而导致点击推特时无法展示发表它的用户。

小组分工

黄孝谨：标签数据绘图，整合绘图，视图交互，制作展示幻灯片

田野：数据爬取与处理，用户数据处理与绘图，视图交互，制作海报

周子淼：绘制推特散点图，视图交互，坐标轴优化，撰写报告

李铄：标签数据处理与绘图，视图交互，时间轴绘制，视频制作