会议总结

2020-8-24

**(彭志远)可视化**

演示了全局视角的octree职业网络以及其上的个人职业层级动态变化。

接下来需要完成个人职业轨迹（trajectory view）的独立展示。

**(刘志威)标签分类**

完成了标签查询系统，根据标签搜索符合要求的候选人员，同时会指定标签是否：

1、满足最低年限

2、处于指定时间段

3、当前仍然任职

后续会对查询结果页面进行优化，显示与标签相关联的关键履历。

**(李可)简历分词**

梳理了简历分词的流程。

**（孔德旭）战疫先锋**

总结了8.20党委组织部会议内容。关于少数分子带动行动力的问题，由任务的完成率、参与率、疫情的控制等可以量化任务的完成效果，从而体现出组织力。

另外用户可能主要在已经存在的社区里传播任务（朋友圈），根据这种任务合作可以推断出，构造一个随机网络，在其上模拟数据传播模型，与实际的模型进行匹配，看能否从合作网络推断出社交网络。

此外讨论了如何进一步结合疫情方面进行分析，从疫情防控的角度提出了两种利用传染病模型，说明自发合作提高组织效能的模型。

**战疫先锋如何结合疫情研究**

**模型1**、使用传染病模型估计出传染率，再与由组织网络的平均距离等指标得到的组织效能进行比较，如因果分析等，来说明自发合作提高组织效能，从而控制疫情的发展。

**模型2**、将深圳市划分为不同的区域，每个区域有独立的传染病模型，区域之间有人口流动，根据防疫任务的时间地点调整相关区域的传染病模型参数（传染率，人口迁移率等），根据不同的组织结构下的仿真结果，说明自发合作能提高组织效能。

区域之间的人口流动的估计方法：从地铁公交/手机定位数据中获取，使用学校或者工厂数量代表。

**问题：**

1、一方面因果关系很难说清，另一方面深圳市内生的病例很少，基本上都是输入型病例以及人口的流动导致，那么疫情爆发时招募志愿者就是资源不足的表现。而排查病情的工作可以预估出人数的需求，因此活动的组织可能是由疫情发展来确定的。所以不能看绝对的总量，而需要对资源和任务进行归一化才能进行比较。

2、小区之间的差异也很大，不同的党支部，可能成员归属感不同，居委会完善一些的话任务也会参与较多，所以需要研究参与自发行动的比例。

3、社会科学难以控制变量，一方面可以归一化，或者减少变量数量；另一方面在传染率不高的情况下，感染人数很少，从而噪声非常大，而相关性分析对于数据的要求很高，所以可以考虑用假设检验的方法，用两类情况进行比较，来说明疫情受到了自发任务的控制。

1. 有病例和没有病例进行比较，关注相应的爆发前的管控措施。
2. 有没有过连锁的感染人数增长的情况（避免感染的偶然性）。
3. 传染链的比较：选出传染链相关的小区，找到出现时间和位置重叠的相关任务。研究传染链是否会因为任务而切断的更快。
4. 比较有且仅有1个病例（没有出现传播）的小区，和超过一例（出现了传播）的小区。对多个小区进行归一化平均。

**意见：**

1. 对比区域或不同时期之间差异后，可以说明自组织对提升组织效能的有效性，从而一方面可以说明算法推荐的best shortcut是有意义的，另一方面，可以在现实防疫工作中推荐优秀的合作任务。在数据中分析发现一些特征，这些特征如何利用到后续的工作中。
2. 可以把对疫情的分析作为亮点，成为组织力的评价之一。