

题目：基于 **Twitter** 的网络结构和社会群体演化

一、问题背景

研究社交互动模式的变化以应对极端事件（地震、海啸等）可以提供对交互行为模式如何响应压力的洞察，可以为规划灾害响应工作提供信息，从而处理例如疏散程序或信息传播等人口层面的行为。

了解极端事件期间用户活动模式的变化需要动态分析个人在极端事件之前和期间与他人交流的方式以及他们的社区交互模式。这些关系均可以映射到网络结构中，其中节点可以是单个用户，边是两个用户之间的交互记录。社区结构可以通过个体群体之间连贯和持续的相互作用模式来观察。

对不良事件下的在线社交媒体的研究是通过利用 **Twitter** 数据开创的。**Twitter** 是一种领先的微博服务，允许用户在一对一、指定的群组或全球基础上，以最多 140 个字符的速度快速交流信息。通过 **Twitter** 观察到的行为并不代表所有的社会互动，但我们相信它确实有潜力展示极端事件如何影响互动模式。

现有一个 **Twitter** 数据集，包含 2011 年日本地震海啸前后用户发文数据，请你们深入分析附件的信息数据，利用数据挖掘的方法研究人们的社会互动模式在受灾害影响下如何改变，并解决下面的问题。其中，附件 1 和附件 2 分别给出了地震前后英文使用者和日文使用者发表的推文数据。

二、解决问题

1.地震前后网络构建

根据附件 1 和附件 2 给出的推文数据，提取出用户发文时间、用户昵称、发文内容等关键字段，并根据这些信息构建地震前后的关系网络。

2.对构建好的网络进行度分析

为了对比地震对用户关系带来的影响，分别对去的英文使用者网络和日文使用者网络进行累积度分布变化统计、单独节点层面的度变化统计并做对比分析。

3.对构建好的网络进行社区检测及对比分析

为了进一步了解地震前后的用户移动情况，请用桑基图显示在每个数据集中地震前后检测到的规模前 10 的社区中的用户移动情况并根据结果做进一步分析。

4.进一步分析

用其他方法，或在时间上进行细粒度动态网络分析。

三、数据说明

附件 1：地震期间英文使用者发帖信息，记录数：7398，格式：json

附件 2：地震期间日文使用者发帖信息，记录数：5450，格式：json

四、参考文献

《Network Structure and Community Evolution on Twitter: Human Behavior Change in Response to the 2011 Japanese Earthquake and Tsunami》