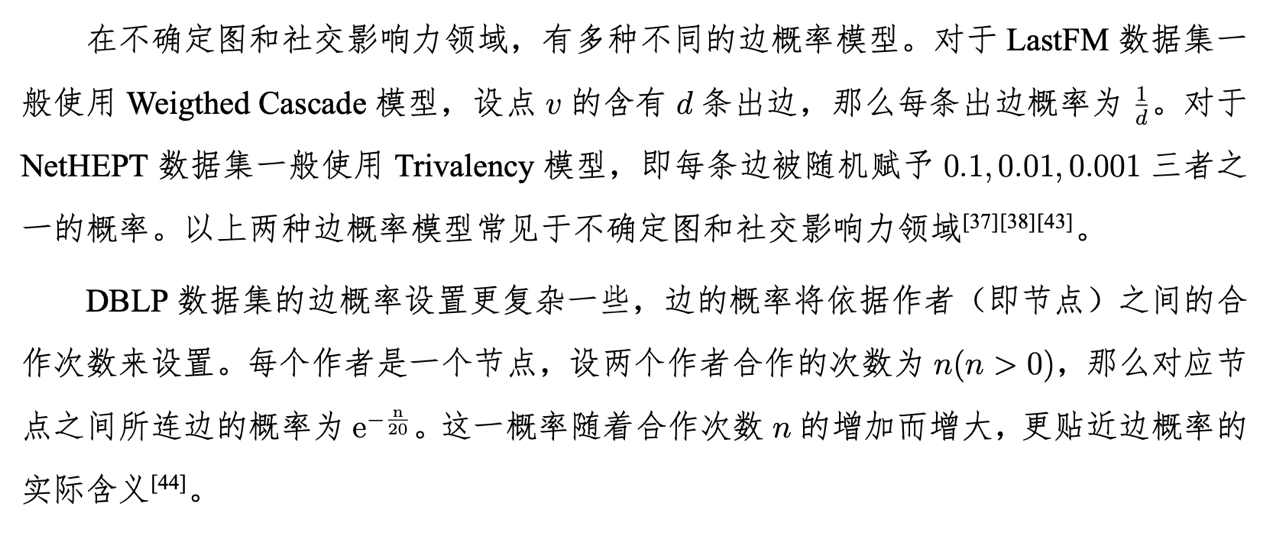
数据集说明

以上数据集均下载自斯坦福公开数据集：<http://snap.stanford.edu/data/>，数据集意义说明在网站上

每个数据集内有两个文件：

* attributes记录了节点数目和边数；
* graph存储了每条边的信息，存储格式为“起点-终点-概率”。其中，概率生成方式为Weighted Cascade模型，具体计算方法为，即终点入度的倒数；对所有小于1e-4的边概率全部处理为1e-4.
* 对无向图的处理方式：每条无向边看成两条有向边，即边看成边和。

除此之外还有以下几种常见的边概率生成方式：



文件夹中部分数据集主要属性如下，如果觉得数据集不合适可以自己去斯坦福网址下载新的数据集；

