

局部社区发现

By: 王浩源, 邝婉静, 谢绮薇, 程文英, 李潇



项目介绍

通过有限的好友资源找出自身处于的社区（一个或多个）

算法：随机游走的标签传播



小组分工

- 搜索模块：谢绮薇
- 标签传播：李潇
- 测试模块：程文英
- 展示 & 网页设计 & 算法设计：王浩源，邝婉静





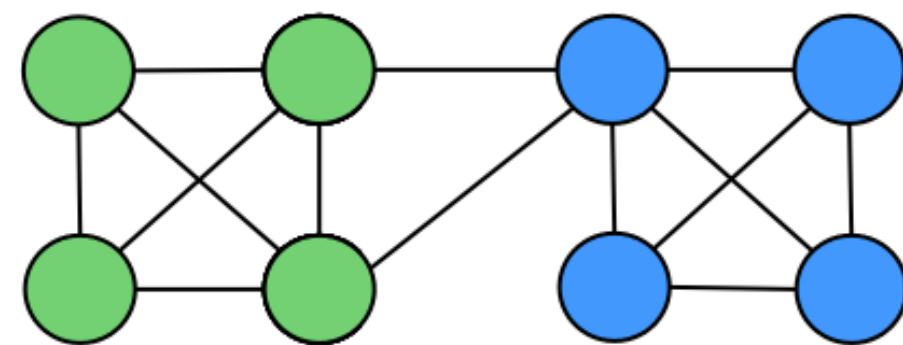
目录

- 提出问题
- 解题思路
- 解题策略：搜索
- 划分算法：标签传播
- 测试
- 存在的问题



提出问题

- 问题：找出一个微博用户所在的社区
- 微博用户之间的关系可以转化为图，节点为用户，边为互相关注关系
- 假设：图中的“团”等于实际存在的社区





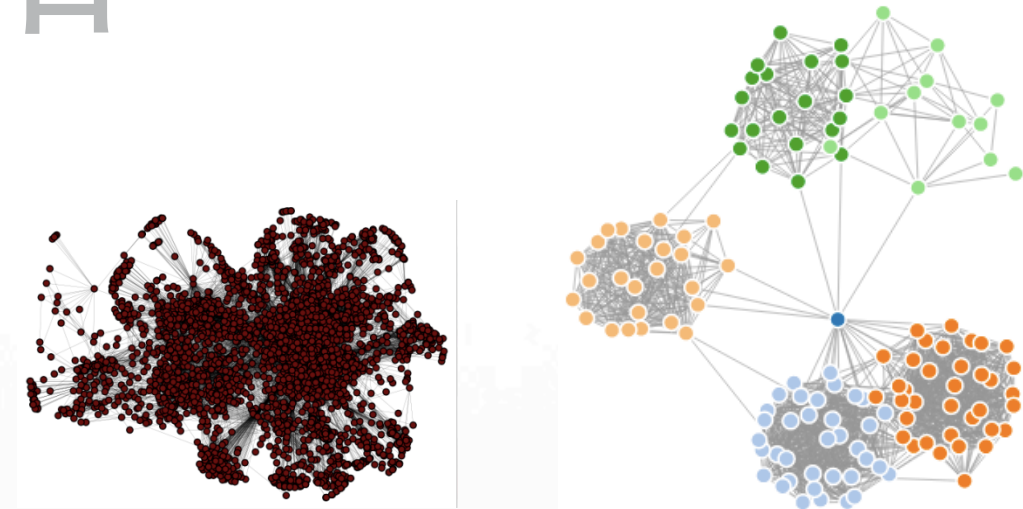
目录

- 提出问题
- 解题思路
- 解题策略：搜索
- 划分算法：标签传播
- 测试
- 存在的问题



问题思路

- 方法一：先搜索全图（两层关系内）
- 社区划分 \rightarrow 子图划分
- 获得全图行不通
- 方法二：搜索中获得社团
- 搜索决策：尽可能直接在搜索过程中找到满足我们目标的点
- 目标：找出与查询用户关联度较大且能体现社区信息的点集





目录

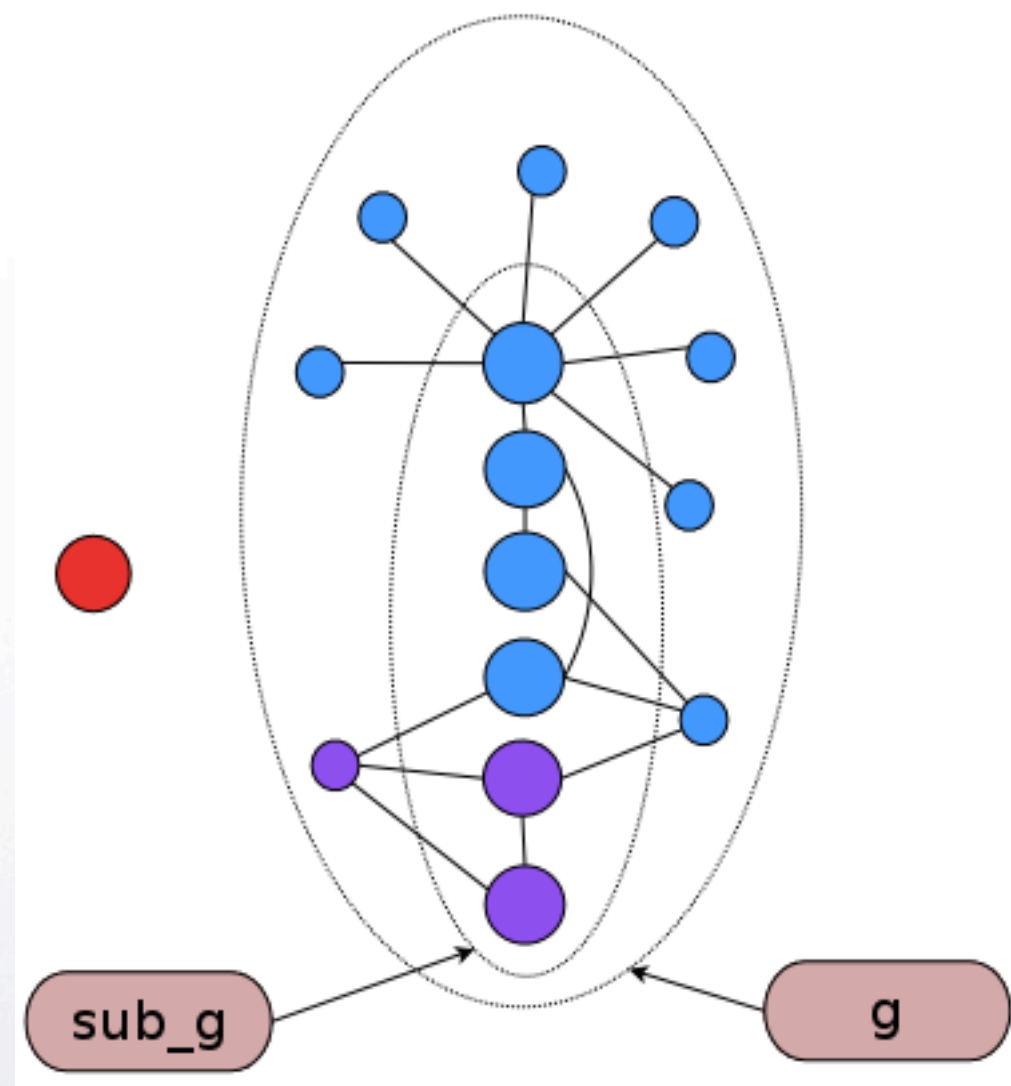
- 提出问题
- 解题思路
- 解题策略：搜索
- 划分算法：标签传播
- 测试
- 存在的问题



搜索模块

- 输入UID
- 展开UID，将除UID外的所有节点加入图中
- 对图进行标签传播划分社区
- 对每一社区中degree高的点展开并加入图中，并重新划分
- 重复至结束

展开：用API得到该用户互相关注列表
标签传播：子图划分方法，后面会详细介绍





目录

- 提出问题
- 解题思路
- 解题策略：搜索
- 划分算法：标签传播
- 测试
- 存在的问题

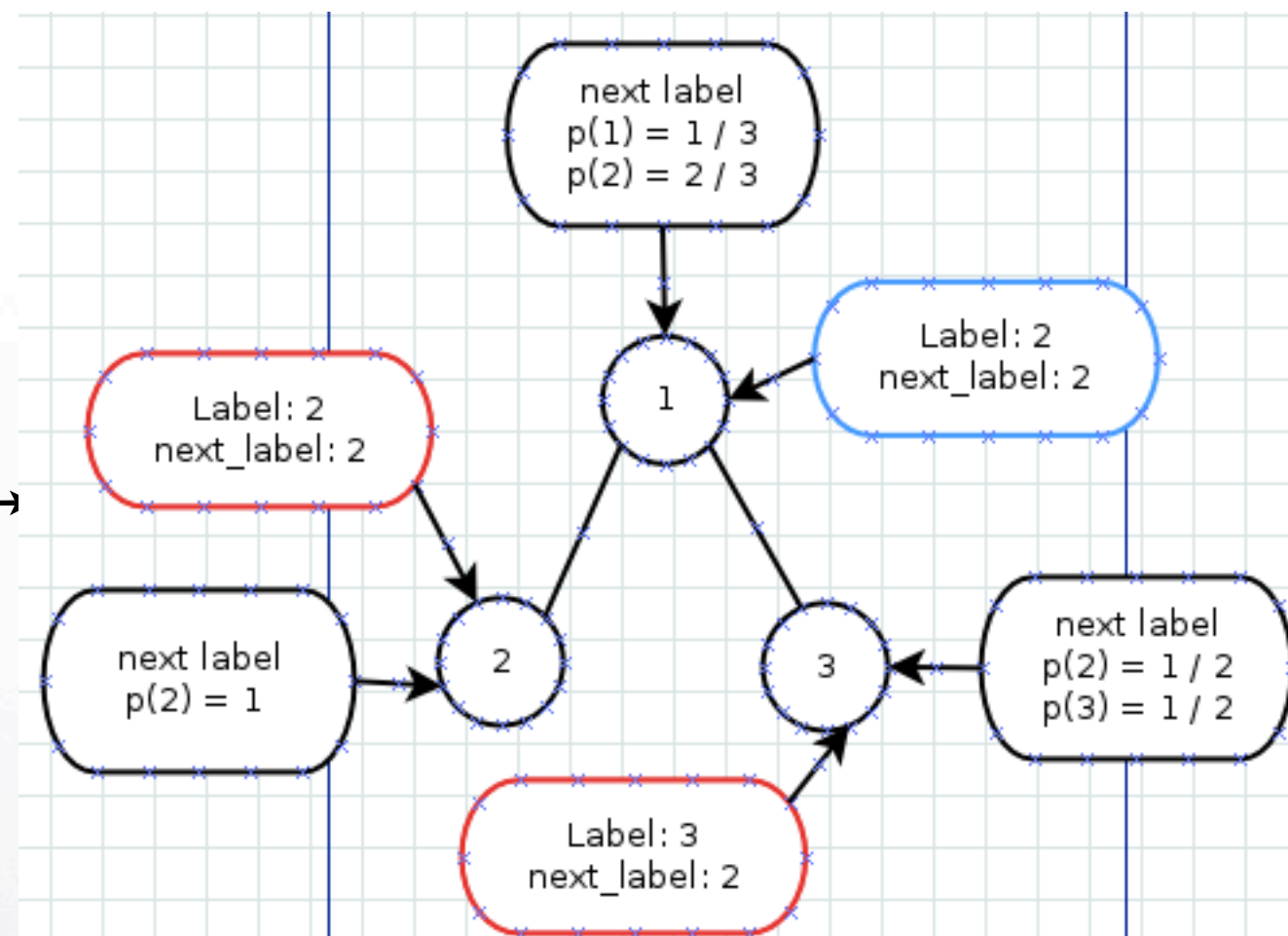
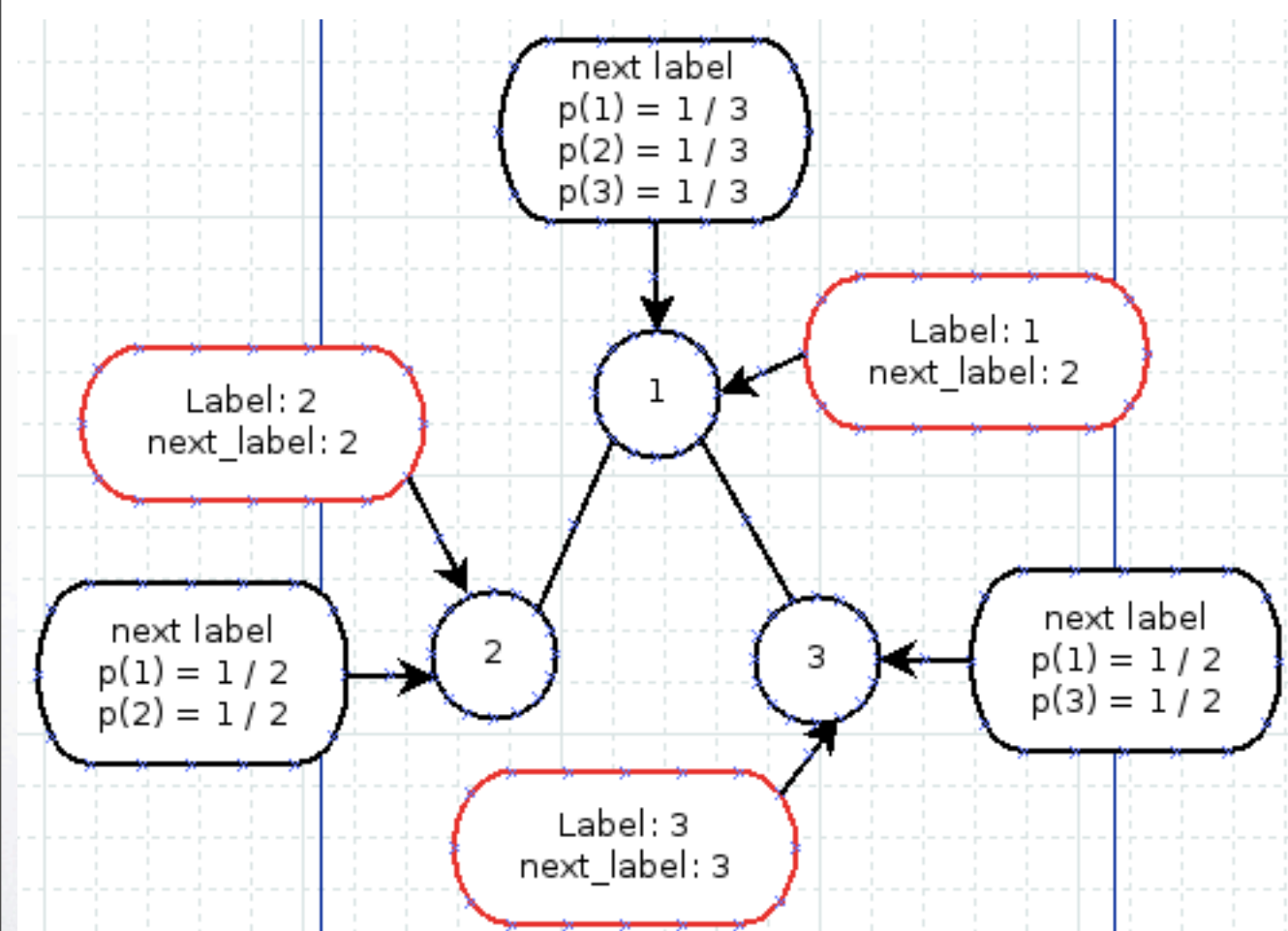


标签传播

- 对一个无权无向图，首先对每一个点赋予一个独一无二的标签（可以理解为属性）
- 每个点选择邻居的标签和自己的标签中出现次数最多的标签为下次转换的标签（次数相同则随即选）
- 全部选择好后集体转变
- 重复传播直至每个点标签不再变化



示意图

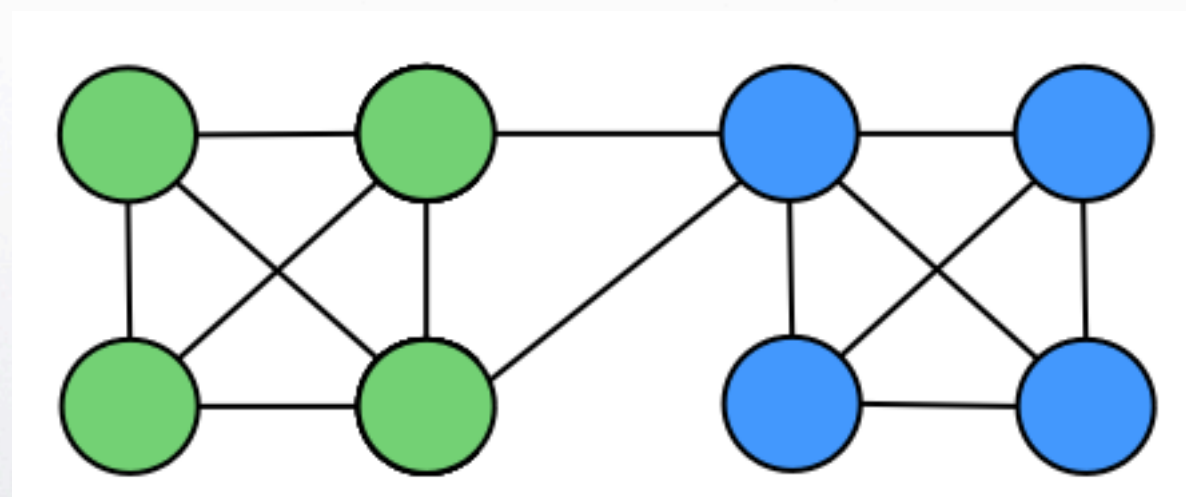
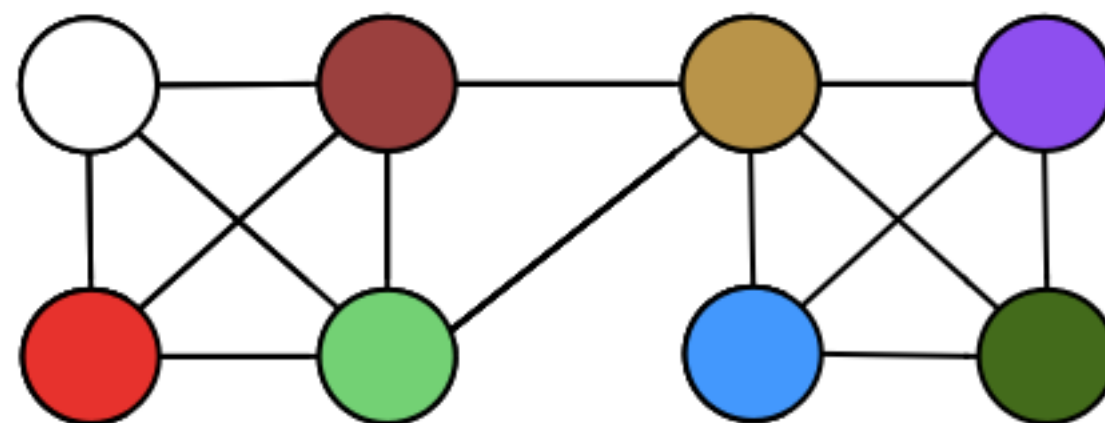
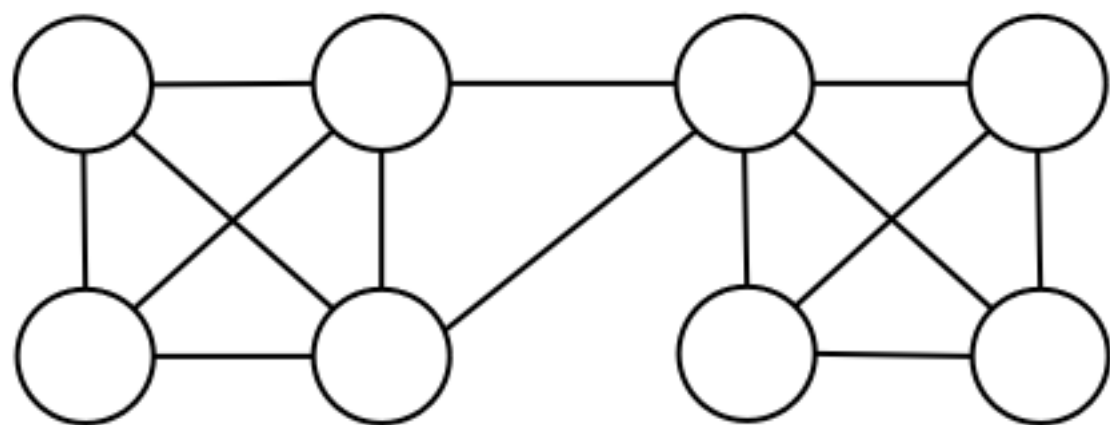


图中红色框内为每个点的属性
黑色框为每一个点的next_label的概率
现在为初始化的示意图

一旦节点1的next_label变成2之后，
便会让2的label固定，几轮迭代后标
签就稳定了

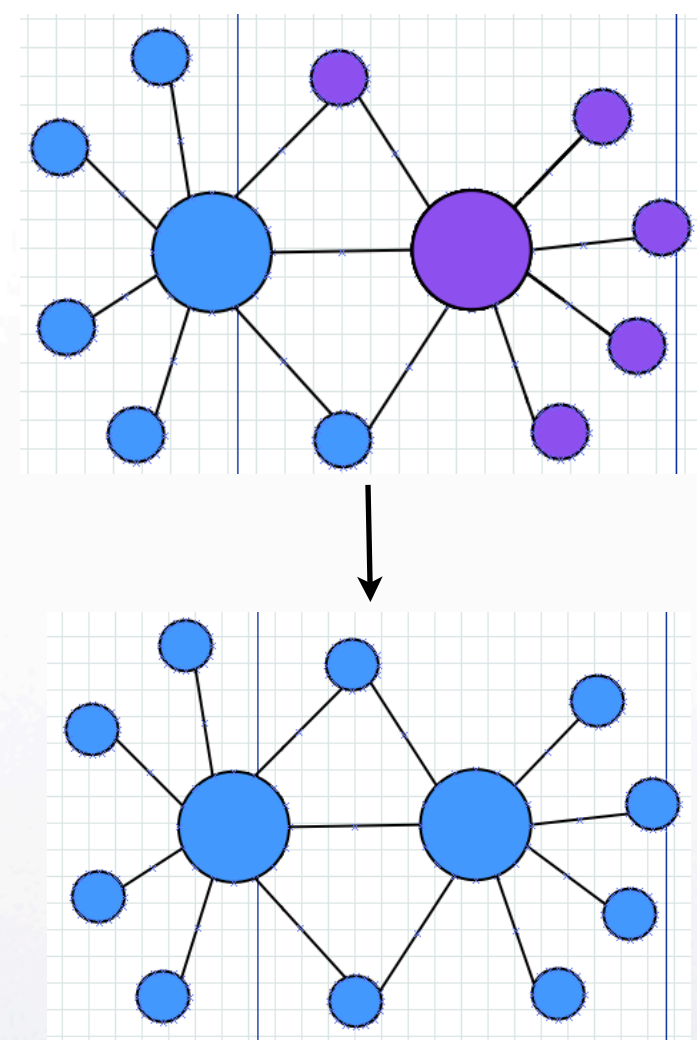
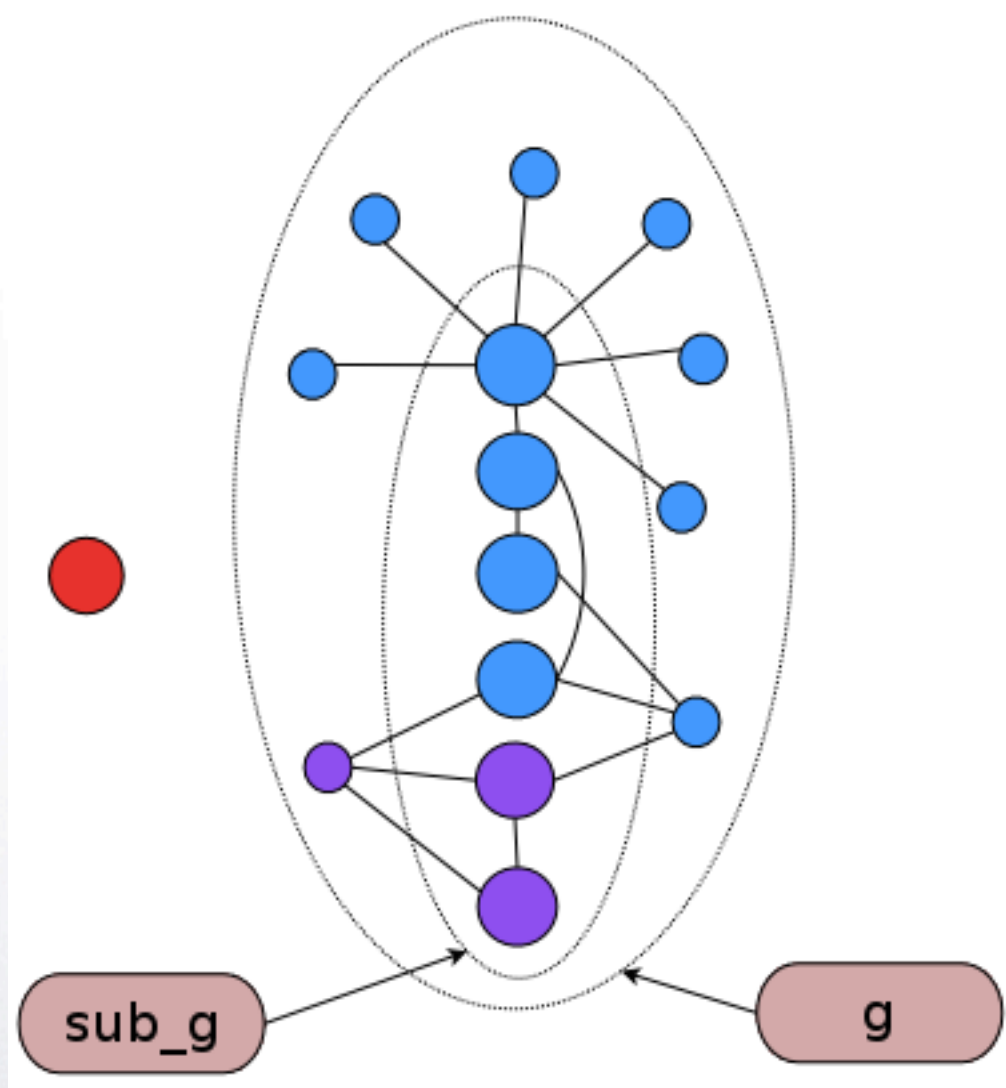


示意图





示意图



大圈为展开过的点



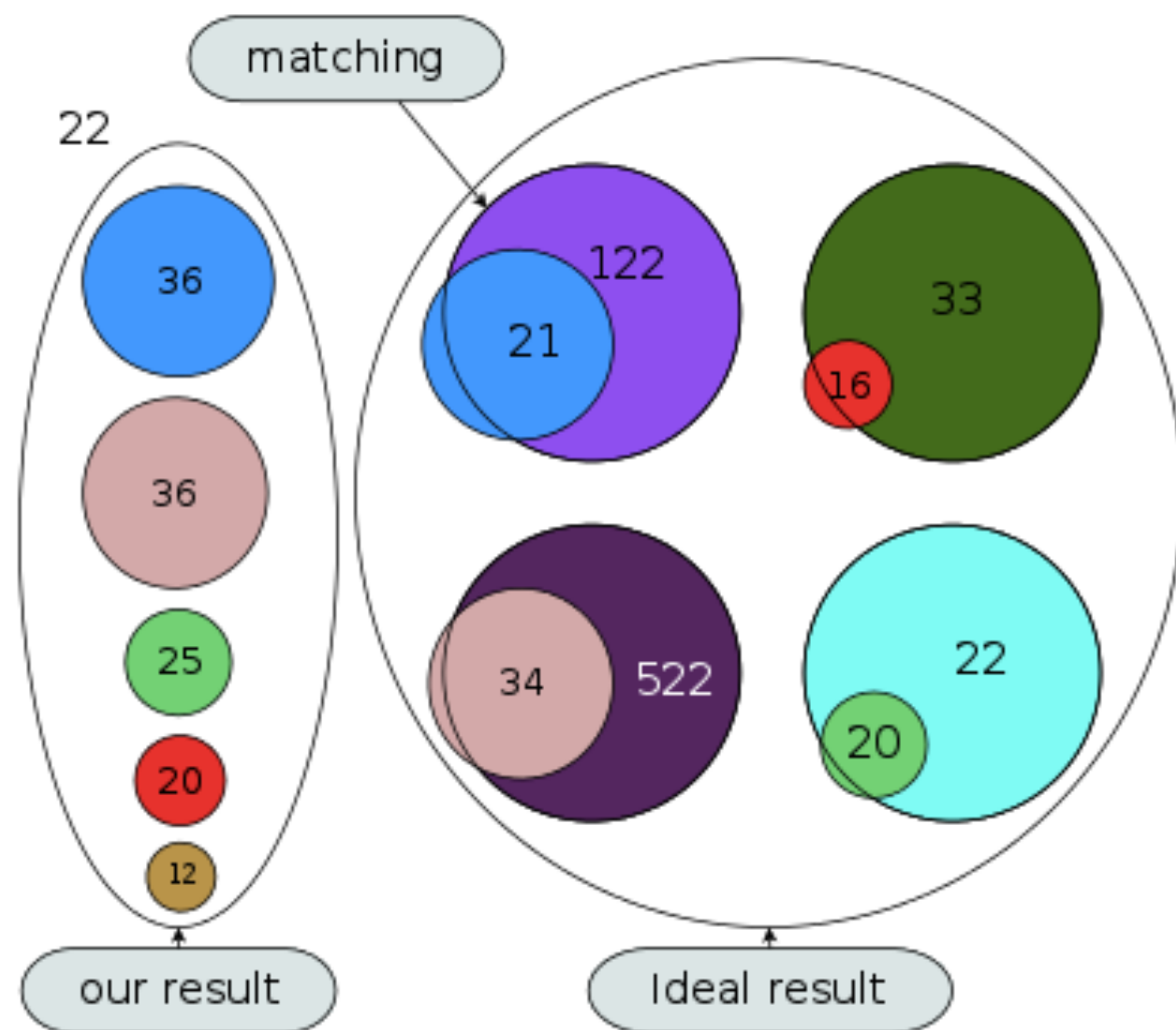
目录

- 提出问题
- 解题思路
- 解题策略： 搜索
- 划分算法： 标签传播
- 测试
- 存在的问题



测试模块

将通过我们的方法划分得到的每一个社区与全图标签传播划分的社区进行匹配，选取最佳匹配的匹配率作为准确度



省略号





示意图





目录

- 提出问题
- 解题思路
- 解题策略： 搜索
- 划分算法： 标签传播
- 测试
- 存在的问题



存在的问题

- 假设：图中的团等于实际存在的圈子
- $\log_2 n$ n 是圈子大小
- 300 每次取的好友列表的最大长度
- $\min(5m, 400)$, m 是直接好友列表长度，整个值是终止条件
- 5 最后保留的社区数上限



结语

谢谢

Thanks for listening
Happy New Year!