

如何对实验结果进行归因:

常规的实验一般步骤:

- 将样本随机划分为实验组与对照组
 - 实验组样本采取某种干预措施, 对照组不采用干预措施
- 根据反馈结果计算, Average treatment effect (ATE)
 - 采用分析两组ATE是否具有显著性差异

一般要满足SUTVA(Stable Unit Treatment Value Assumption)假设

- 实验组与对照组之间不存在溢出效应
- 实验组的行为不会影响对照组样本的行为



社交网络实验的独特性:

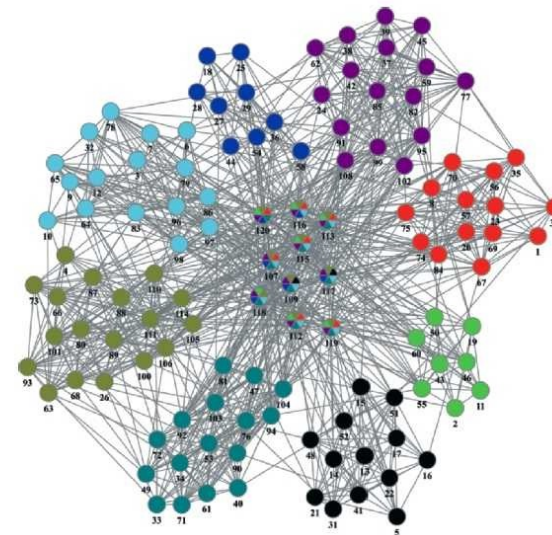
社交网络实验一般很难满足SUTVA假设

- 无法完全消除网络边的溢出效应
 - 社交网络的中玩家之间存在社交关系，这种社交关系会影响实验组与对照组的行为决策，进而导致ATE值无法反映客观想象

社交网络实验一般要尽可能消除或减轻网络效应的影响

- 在流量切分时，尽可能把社交好友划分到同一个流量内
- 不同流量之间玩家边的连接尽可能稀疏

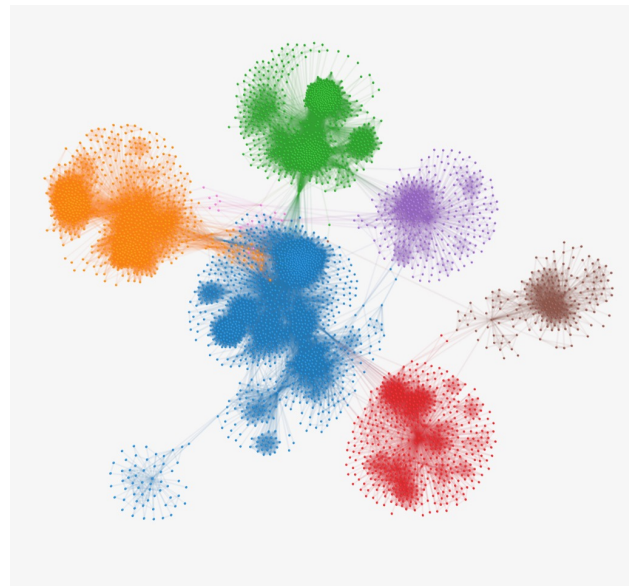
- 可以利用社群划分结果进行流量切分



流量内的玩家边连接尽可能稠密，
流量间玩家的边连接尽可能稀疏

采用社群划分结果进行流量切分存在的问题:

- 社群大小不均匀，可能导致流量切分并不一定成比例
 - 大社群可能会带来更大的流量，进而破坏流量的平衡
- Fastunfolding算法没有考虑节点的特征
 - 单纯依据好友关系链进行划分社群，可能导致特征分布不均衡
 - 在付费、活跃度、好友数等指标上，不同流量样本之间可能会存在差异



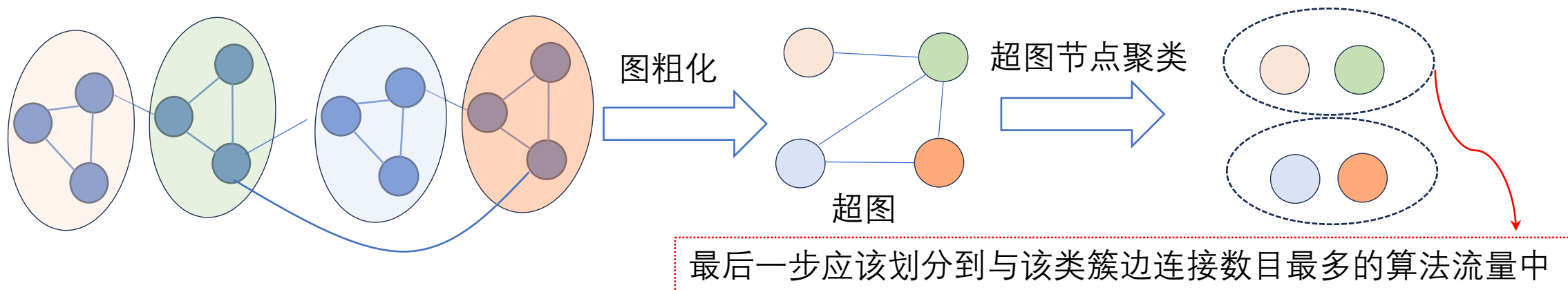
网络效应实验的研发方案:

- FastUnfolding算法, 可以把网络划分为若干个社群, 以减小边的溢出效应

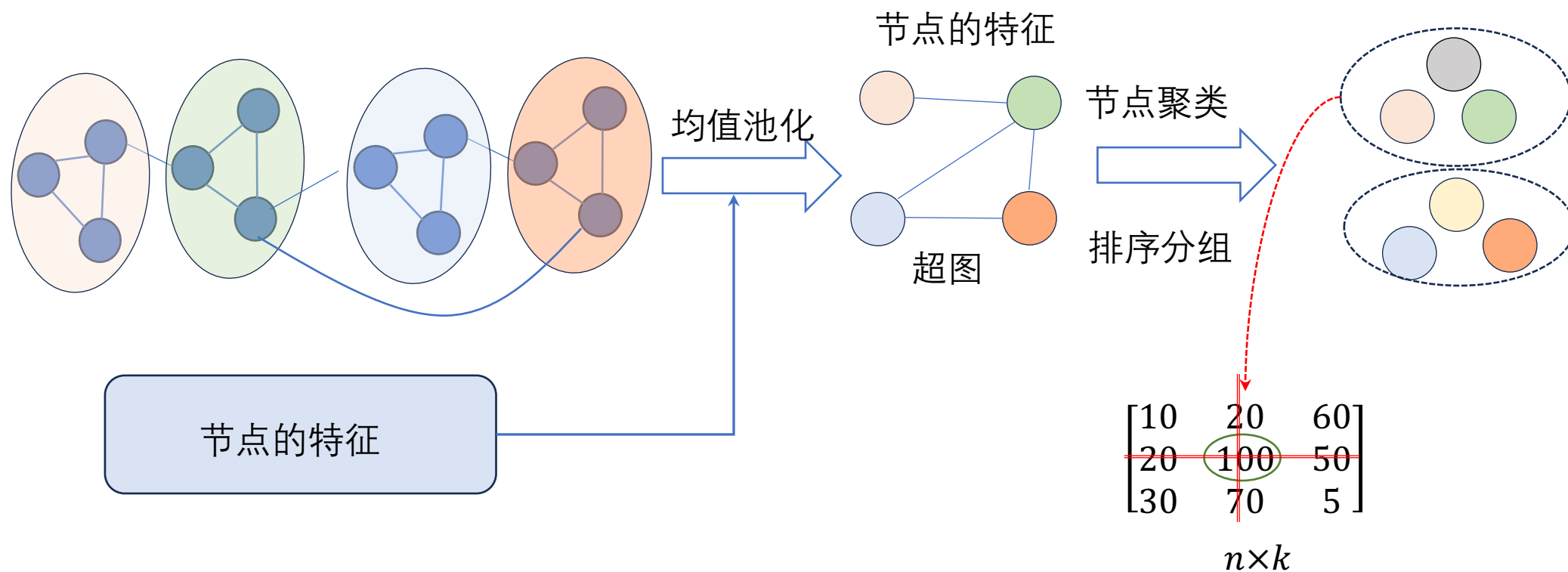


如何保持节点特征的均匀性?

采用图粗化(graph Coarsening)方法保持, 将子图匹配问题, 转化为超图节点匹配问题



超图节点如何匹配？



- 同一类簇内， k 个社群为一组，按照与流量内社群最大连接性依次进行分配

实验对比方案:

- 以XXXX好友网络为实验数据，全量玩家切分为4份均等流量
- 对比方法
 - 完全随机: openid+roleid随机映射
 - 基于社群结构的流量切分: 社群ID随机映射

网络效应实验的评价指标:

边的纯度=类簇间的边数/图的边数

- 体现了消除溢出效应能力，越小越好

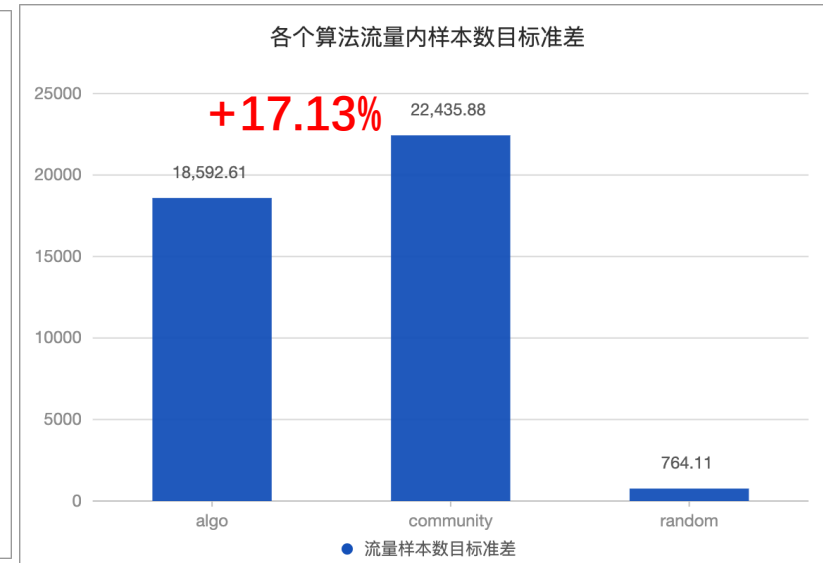
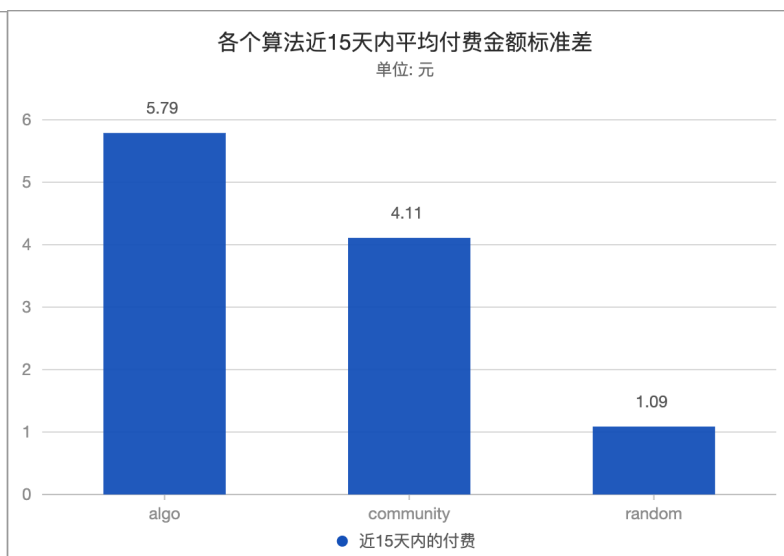
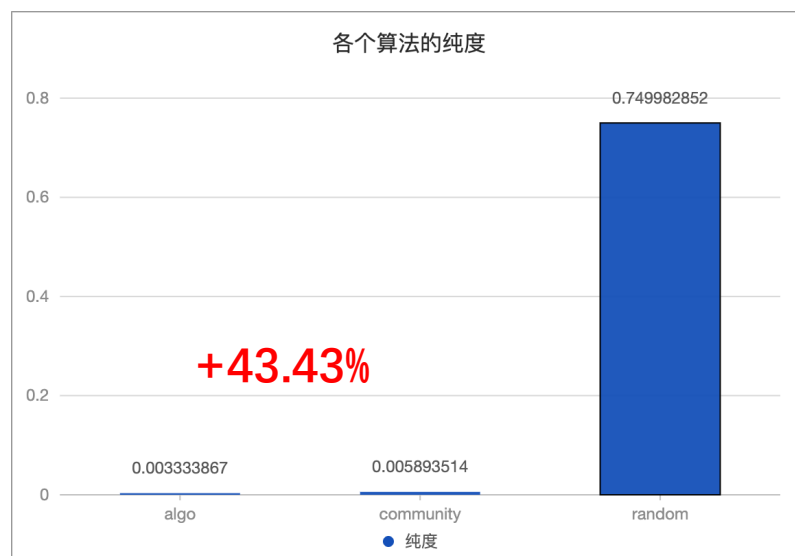
流量内样本数目标标准差

- 标准差的大小体现了节点在流量内的聚集程度，越小越好

各个流量下的相关特征指标的均值标准差

- 近15天内各个流量付费均值标准差，越小越好

实验结果(15天实验结果的均值)



- 相对于社群ID映射方法，采用新算法可以进一步降低边的溢出效应，并获得更加均匀的流量切分结果
- 节点的付费特征选择存在不合理之处，每天的付费特征难以反映人的付费能力，应采用付费数据的统计特征更合理

理想的流量切分方案具有的特点:

- 无网络效应

- 各个流量样本之间边的溢出效应要尽可能弱



- 节点数目分布均匀

- 同等流量下样本数目差异要尽可能小



- 特征分布均匀

- 在付费、活跃度、好友数等指标上，不同流量样本之间差异要尽可能小



- 拓扑结构分布均匀



- 流量间边的连接数目要均匀
- 流量内边的连接数目要均匀
- 流量划分后图的拓扑结构要分布均匀(节点度的分布)

社群大小不均匀
社群节点特征不均匀

现有方案的优化:

- 找一个平衡性图聚类算法, 满足子图大小均匀, 并且满足多个约束条件
- reLDG算法可满足平衡性聚类的问题:

$$\arg \max_{i \in 1 \dots k} |C_i \cap N(u)| \left(1 - \frac{|C_i|}{C}\right)$$

$N(u)$ 为节点 u 的邻居节点集, C_i 为第 i 个类簇, $C = \frac{n}{k}$ 为类簇的最大容量



加入条件约束

$$\arg \max_{i=1,2,\dots,k} |C_i \cap N(u)| \left(1 - \frac{C_i}{C}\right) \prod_l (\bar{y}_{i,l} - \bar{y}_l)(\bar{y}_l - y_{u,l})$$

q分-邻域的节点才会纳入ATE计算

$\bar{y}_{i,l}$ 为第 i 个类簇在 l 特征上的均值, \bar{y}_l 为全体对象在 l 特征上的均值, $y_{u,l}$ 为节点 u 在 l 特征上的取值

Nishimura J, Ugander J. Restreaming graph partitioning: simple versatile algorithms for advanced balancing. KDD 2013: 1106-1114.

Saveski M, Pouget-Abadie J, Saint-Jacques G, et al. Detecting network effects: Randomizing over randomized experiments. KDD 2017: 1027-1035.

Shang H, Shi X, Jiang B. Network A/B Testing: Nonparametric Statistical Significance Test Based on Cluster-Level Permutation. Journal of Data Science, 2023, 21(3):523-537.

Thank you