

H1 Adaptive Seeding in Social Networks

H2 背景

在社交网络中，最大化以口碑方式的信息传播被广泛的研究，已经有了很多的算法框架和技术库。最为经典的是IM问题，此后又延伸出了许多问题。但这些研究都对网络访问做了理想化的假设，即可以访问或者选择网络中的任何节点，而实际上要得到全局信息并不轻松。由于网络节点是根据度数呈重尾分布的，即度数高的节点不容易被观察到，那么可以考虑将IM应用到样本领域中，因为高度节点有很多邻居，人们希望这样的节点将被样本连接。

因此，尽管大多数节点在信息传播中不起作用，但它们有播种其邻居节点的潜力。据此，作者提出了一个二阶段播种过程，第一阶段通过部分预算选择样本节点，并邀请它们的邻居加入可访问集合，第二阶段，当一些可访问邻居被播种以后，剩余预算被用来从可访问集合中选择另一部分影响者。

H2 解决问题

- 解决了实际中对社交网络访问存在限制而不能有效地应用现有技术解决IM的问题。

H2 创新之处

- 为影响最大化研究提供了另一种思路，在网络限制访问的情况下实现影响最大化的研究。
- 开发了一种二阶段随机优化框架Adaptive Seeding用来解决这个问题。

H2 解决方法

H3 Adaptive Seeding

H4 Define

Adaptive Seeding是一个二阶段过程，第一个阶段初始化可被选择节点集合 $X \subseteq V$ ，第二阶段每个被选择节点的邻居以独立概率 p_j 出现，并可以被选择。对任何邻居集合 R ，用 p_R 表示被实现的可能，那么 $P_R = \prod_{i \in R} p_i \prod_{i \notin R} (1 - p_i)$ 。给定一个预算 k ，自适应策略的目标是选择 X 的子集 S ，其大小 $t < k$ ，在剩余预算下影响函数 $f: 2^V \rightarrow \mathbb{R}_+$ 在 S 所有可能实现的邻居上最大化。用 $\mathcal{N}(Q)$ 表示 $Q \subseteq S$ 的邻居，对于任意 $S \subseteq X$ ，让 $S_R = \operatorname{argmax}\{f(T) : T \subseteq R \cap \mathcal{N}(S), |S| + |T| \leq k\}$ 。最优解为：

$$S^* := \operatorname{argmax}_{S \subseteq X} \left\{ \sum_R p_R \cdot f(S_R) \right\}$$

H4 Function

影响函数 $f(S)$ 的值是在一个随机过程中，初始化种子 S 影响其它节点的期望数量。

H4 Randomized-and-relaxed policies

文章主要技术涉及到一个随机松弛非自适应策略，它提供主要思路。一个策略是非自适应的如果它只出现在第一阶段，第二阶段在节点实现之前，同时提交 S 和 Q ，它将会进行选择。这样的策略返回了一个集合 S 和一个权值 q_i ， q_i 是节点 i 被实现后，被选择的概率。该策略最大化问题为：

$$\begin{aligned} \max \quad & \sum_{T \subseteq \mathcal{N}(S)} \left(\prod_{i \in T} p_i q_i \prod_{i \notin T} (1 - p_i q_i) \right) f(T) \\ \text{s.t.} \quad & S \subseteq X \\ & |S| + \sum_{i \in \mathcal{N}(S)} p_i q_i \leq k \end{aligned}$$

H4 Step

1. 构造非自适应随机松弛问题的凹函数作为优化目标
2. 计算出最优自适应解和非自适应解相差的常数因子，从而用非自适应策略解估计适应策略解
3. 通过爬山算法在给定的整数和分数限制条件下求解目标的近似解，每次迭代算法将会模拟梯度上升的过程求解。

H3 Simple Approches

如果影响函数满足单调性和子模性，可以通过采样技术去近似求二阶段问题的最优解。

分两种特殊情况：

- 可加函数

对于任意 $S \subseteq X$ ，选取在 $\mathcal{N}(S)$ 中函数值最高的节点，添加 t 个最高值的节点到 S 中，返回 S 。

- 对称子模函数

加入新元素到一个更大的集合会引入较少的邻居，每个新邻居节点的边际贡献会随着邻居节点的增长而减小。

H2 相关工作

- Asadpour et al. and Krause 研究多阶段变体影响最大化问题，种子被序列化选择，但在每个阶段，整个节点集合是可用的。
- Hartline et al. and Haghpahanah 考虑数字商品的利益最大化，个人对某项商品的估价是其邻居采用该商品的函数。
- 其它还有在随机优化问题上做的研究，其中 Dean et al. 研究自适应情境，并定义了自适应间隔概念，他们构造了非自适应策略完成了最佳自适应的值的恒定比例。

H2 想法

该文章为网络限制访问时影响最大化的研究提供了一个新思路，同时给予了一定的理论支持。虽然将其称为自适应（渐进式），但实际只有两个阶段。整篇文章有些晦涩难懂，大概看懂了它的思想，但因为文章没有实验，不能直观地去理解，也不能看到其效果。