作业4：影响最大化和信息级联

- 信息级联模型的简单贪心算法：
- 输入：图$G=(V,E)$，k

步骤1：种子集合$S=\{\}$;

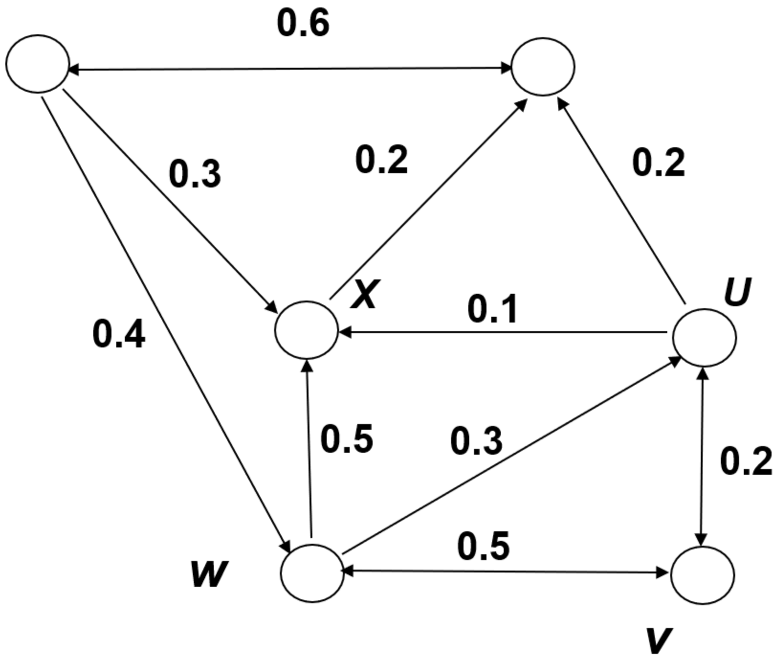
步骤2：若S的大小达到k，直接返回S;

步骤3：对于所有的候选节点$v \in V \setminus S$，计算v的边际收益$R(v \cup S)-R(S)$，这里的回报即是激活节点数;

步骤4：选取使边际收益最大的节点v加入S，回到步骤2.

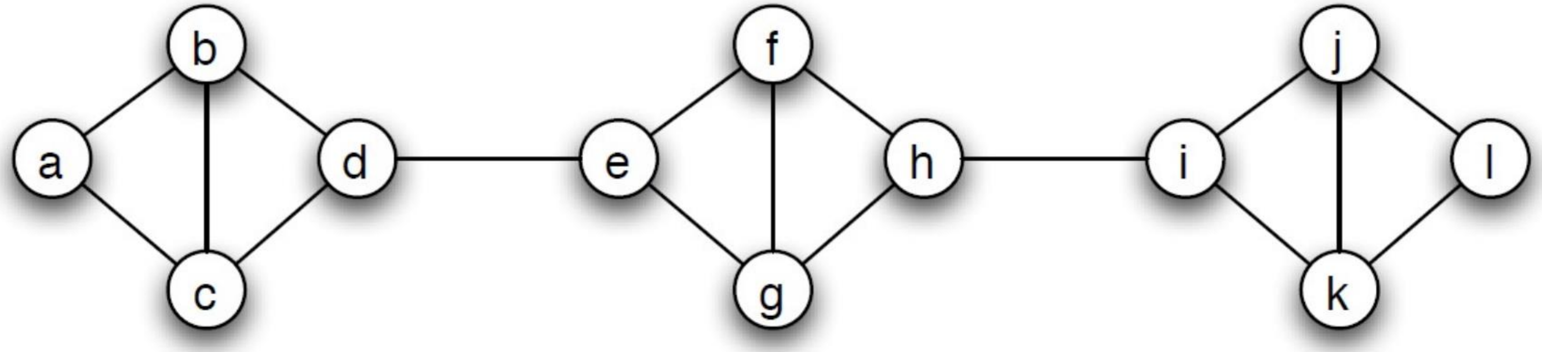
作业4：影响最大化和信息级联

- Implement simple greedy algorithm for influence maximation with independent cascade model;
- Test the algorithm on the following graph (10 points):



可选题目：信息级联定理

- 恒定阈值信息级联模型（基于P15的Linear Threshold Model)：
  - 每个节点存在一个恒定的阈值$q(0 < q < 1)$，每一次迭代每个节点邻居被激活的比例大于等于q时，激活该节点；
- 簇(cluster)的定义：
  - 称一个节点集S为密度p的cluster，若S每个节点至少有比例p的邻居属于S
  - 簇的例子：3组4个顶点，密度为2/3的簇。



可选题目：信息级联定理

- 可选题-证明（信息级联定理）：
- 设在图G中S是初始的激活节点集，且剩余节点的激活阈值是q

  1. 若剩余网络包含一个密度大于$1-q$的簇，则G不能被完全激活（迭代收敛后，G中仍然有节点未被激活）；

  2. 若G不能被完全激活，则剩余网络必然包含一个密度大于$1-q$的簇。
- 这个定理与社会网络的哪些性质有联系?