

社会网络

1. betweenness (中介度): 经过该边的最短路径数 (多条最短路径时按比例贡献).

2. Girvan-Newman 算法: \rightarrow 计算所有 betweenness, 每次删除最大一条边.

两个问题

- ① 怎样计算 betweenness
- ② 怎样选择 clusters 数目.

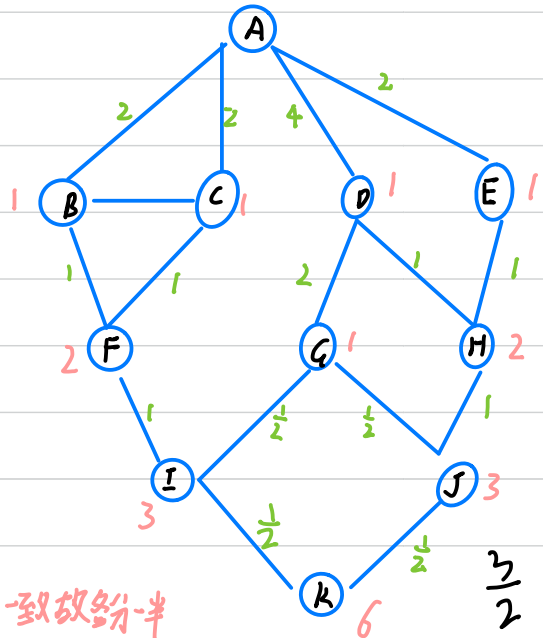
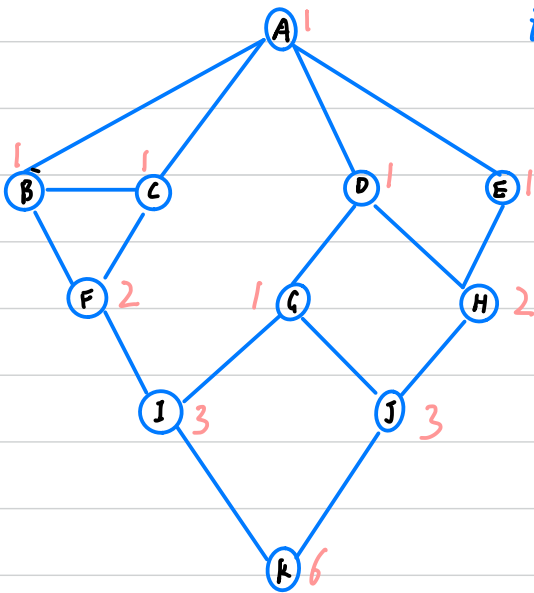
1. 将原图形状拽成 tree.

计算A到各个节点的最短路径.

e.g.: shortest A-J = shortest A-G + shortest A-H

故节点如图所示.

J: $\frac{1}{2} + 1$, G和H分别占 $\frac{1}{3}$ 和 $\frac{2}{3}$



2. 计算 betweenness:

node flow = $1 + \sum$ 每个 child 提供的边.

该节点向上面父节点的边分配边是根据

父节点的 shortest 占比分配.

K: $1+0$, I和J shortest 一致故各分一半

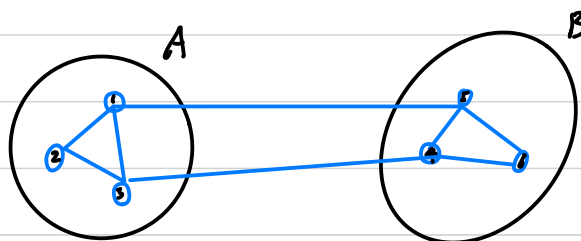
$\frac{3}{2}$

将每个节点都当根节点，计算一遍后，将每条边分值求和并除以2
(无向图计算2次)。

群落(communities): 紧密相连的节点集合。

3. spectral clustering — 谱聚类

cut: 两个集合中各取一点，能组成的边数量。



$$\text{cut}(A, B) = 2$$

归一化切割: Normalized-cut

$$\text{ncut}(A, B) = \frac{\text{cut}(A, B)}{\text{vol}(A)} + \frac{\text{cut}(A, B)}{\text{vol}(B)}$$

$\text{vol}(A)$: 至少有一个端点在A的边总条数。

$$\text{vol}(A) = \sum_{i \in A} W_{ij}$$