

社交网中同质性的一种 测量方法

社交网中的同质性现象

朋友～相似

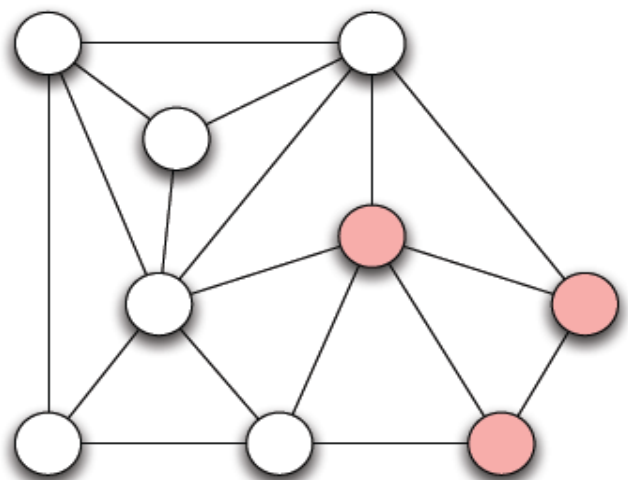


- “相似”的含义可因考虑的问题不同而不同
 - 固定特征，可变特征

如何定量评估一个社交网络中同质性现象的程度？

- 给定社交网（只考虑两种不同特征：红，白）
- 我们能得到的信息
 - 节点数 (n)，边数 (e)
 - 不同颜色节点的占比： p ， $q=1-p$
 - 两端节点相同的边数 (s)

基于它们，如何讨论同质性现象的程度？



节点数 $n = 9$

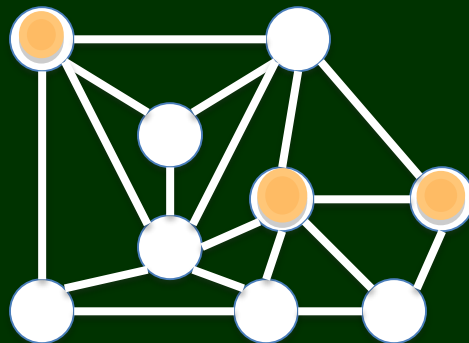
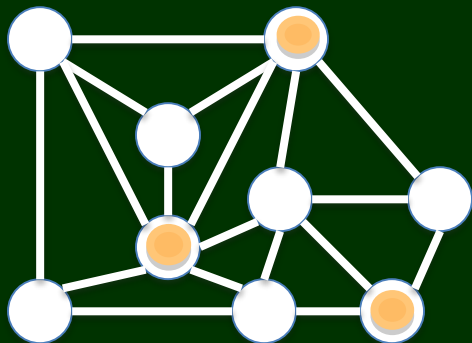
边数 $e = 18$

红色节点占比 $p = 1/3$

白色节点占比 $q = 2/3$

两端节点相同的边数 $s = 13$

- 认识：两端节点相同的边越多（占比越高），同质性越明显



- 认识：两端节点相同的边越多（占比越高），同质性越明显

$$\frac{s}{e} = ?$$

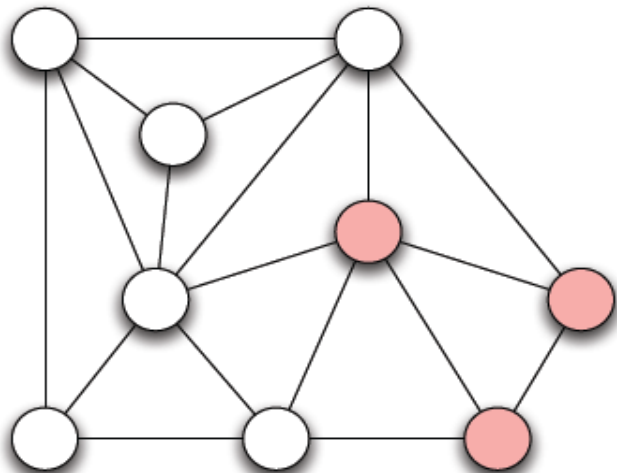
有没有一个基准？多少就算多？

用“随机情况”作为基准：给定不同颜色节点的占比（红 p 和白 q ），随机情况下，一个节点是红色的概率就是 p ，白色的概率就是 q ，那么任何一条边的两节点颜色相同的概率就是 p^2+q^2 ，也就是两端节点相同边的占比。



$$\frac{s}{e} > p^2 + q^2 ?$$

- 节点数 (n) , 边数 (e)
- 不同颜色节点的占比: p, q=1-p
- 两端节点相同的边数 (s)



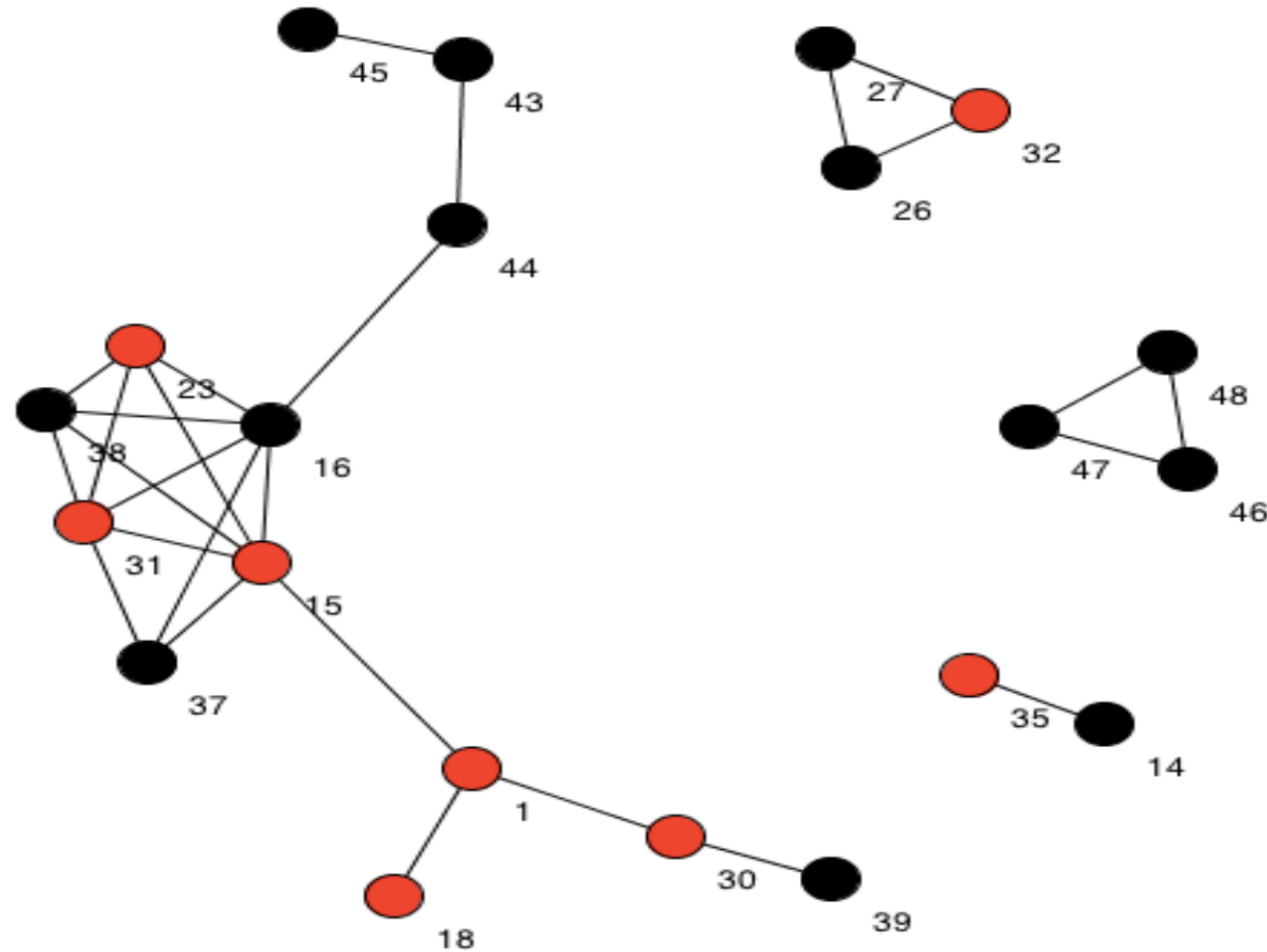
$$\begin{aligned} n &= 9 \\ e &= 18 \\ p &= 1/3 \\ q &= 2/3 \\ s &= 13 \end{aligned}$$

$$\begin{aligned} s/e &= 13/18; \\ p^2 + q^2 &= (1/3)^2 + (2/3)^2 \\ &= 5/9 = 10/18; \end{aligned}$$

因为两端颜色相同的边数的占比
大于作为基准的平均情况,
所以我们可以得出下面的结论:

同质性现象在这个
社交网中有所表现

Quizz:



此图源于一个真实班级在最初的社会网络，其中有21个节点，27条边。红色代表女生（8名），黑色代表男生（13名）。两端点相同的边有15条。试判断此网络是否显现同质性。

小结

- 社交网络中的同质性现象有可能定量表达，这一节介绍了一种最基本的方式，其精神也可以被推广到其他更复杂的情形
- 在这一节的讨论中，我们采用了和教材中稍有不同、但是等价的讲法（参见教材 4.1节）