

# 同质性：影响社会网络结构的重要外部因素

前面讨论的“三元闭包”影响网络结构，属于网络自身的因素（内部因素）

# 提要

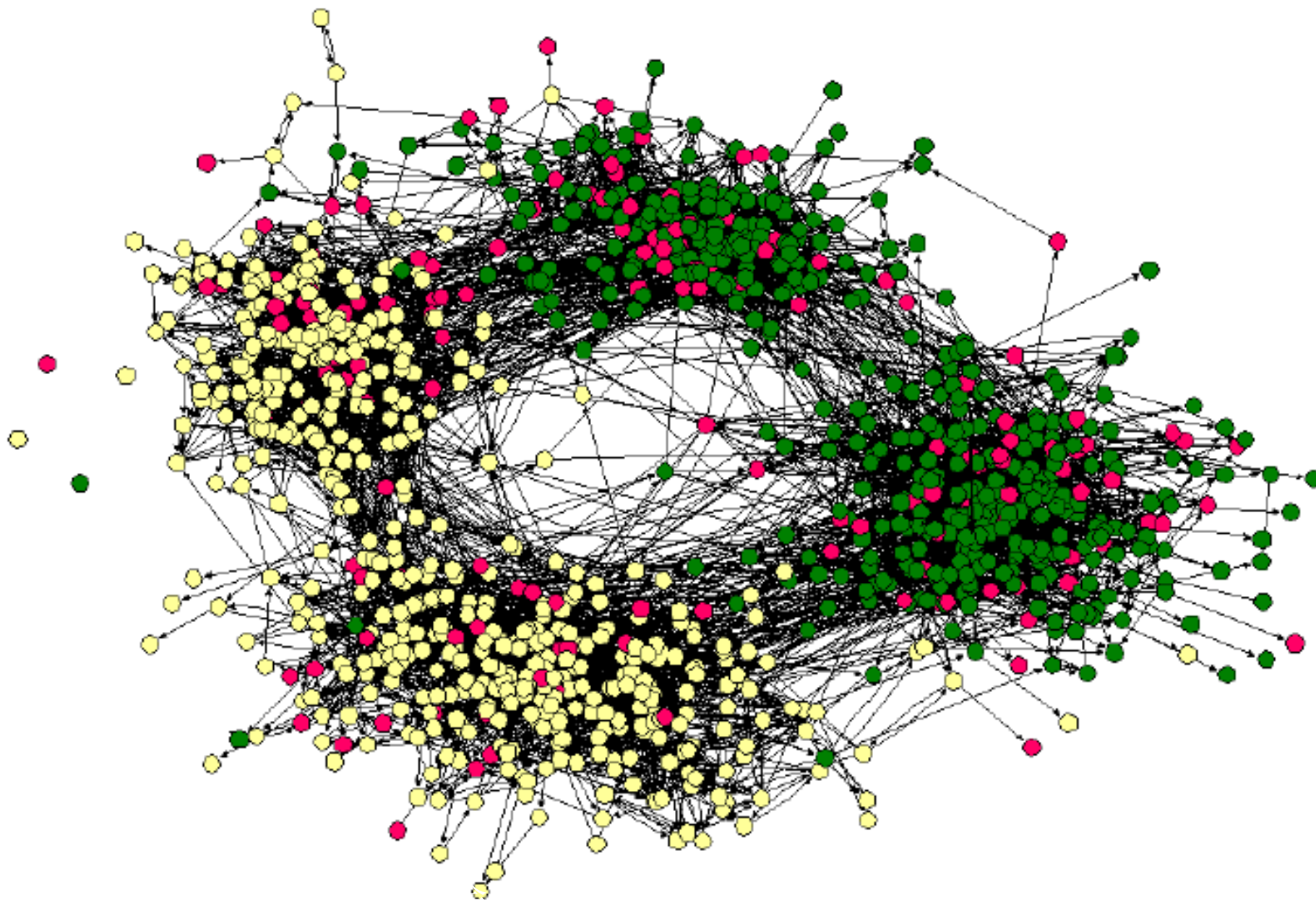
- 同质性的基本含义
- 同质性在网络中的一种评估方式
- 归属网、社会归属网对同质性的影响
- OSN数据分析例子
- 隔离——同质性的一种现象，谢林模型

# 同质性 (homophily)

物以类聚，人以群分  
近朱者赤，近墨者黑

- 每个人的特质可分为两种：
  - 固有特质：性别、种族、母语等
  - 可变特质：居住区、爱好、专长、思想等
- 基本问题：因为相似，才成为朋友（**selection**）；还是因为成为朋友后变得相似（**social influence**）？
  - 给定相似性，区别“选择”与“社会影响”的作用程度是社会学研究中一个经典问题

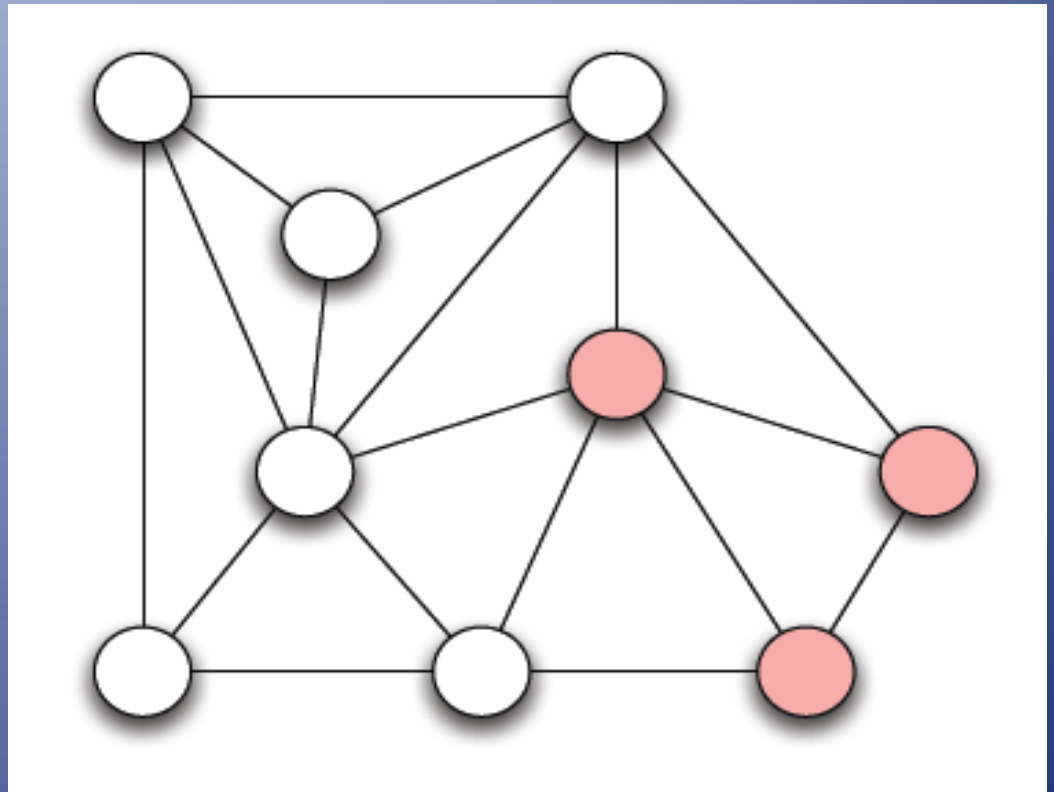
# 一个反映较强同质性的社会网络



# 社会网络中同质性判别的一个测度

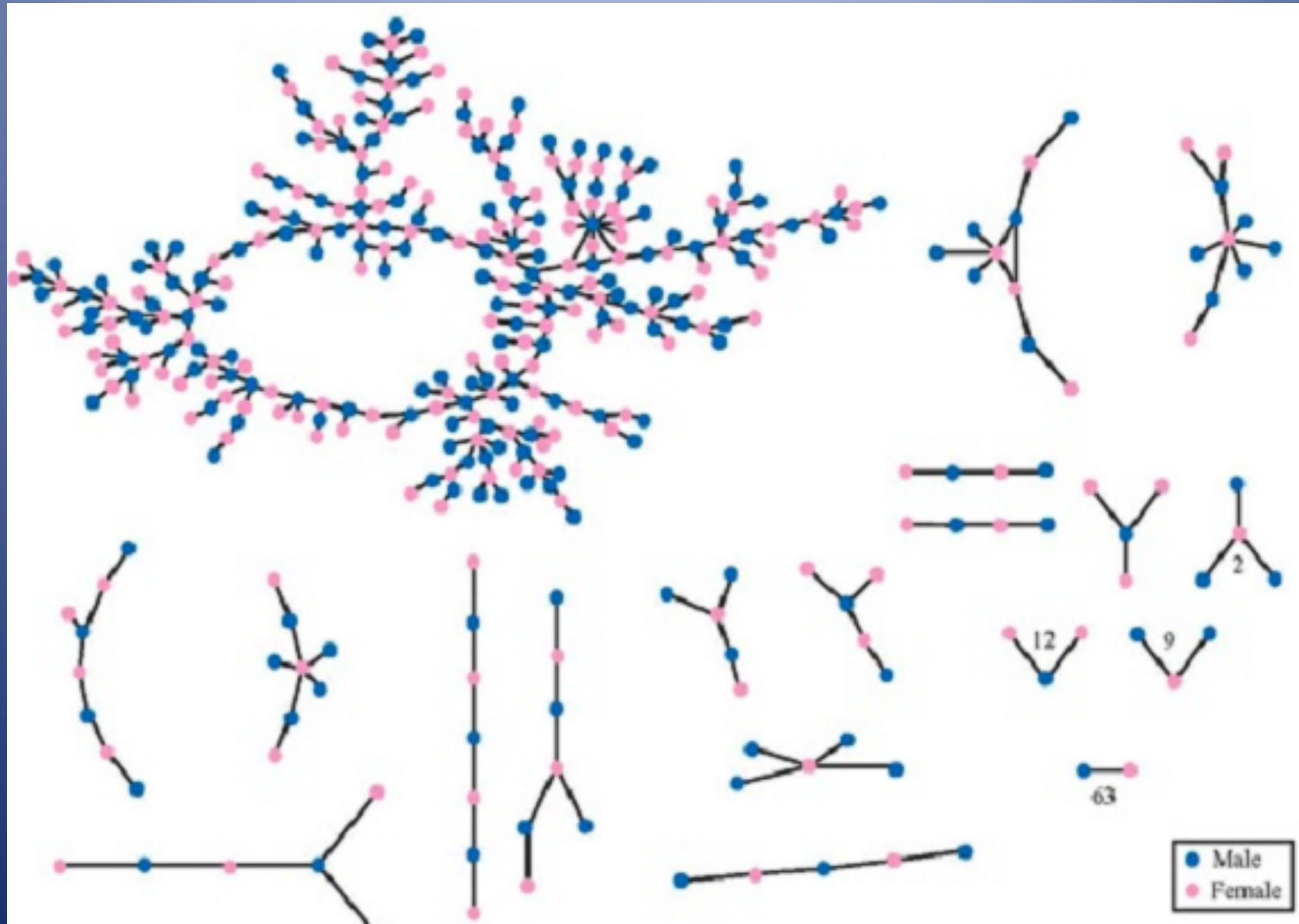
- 假设人们按照两种属性区分（例如性别），我们来看如何判断社会网络中同质性体现的程度

- 直观上：相同颜色的节点的聚集程度
- 如何定量把握？



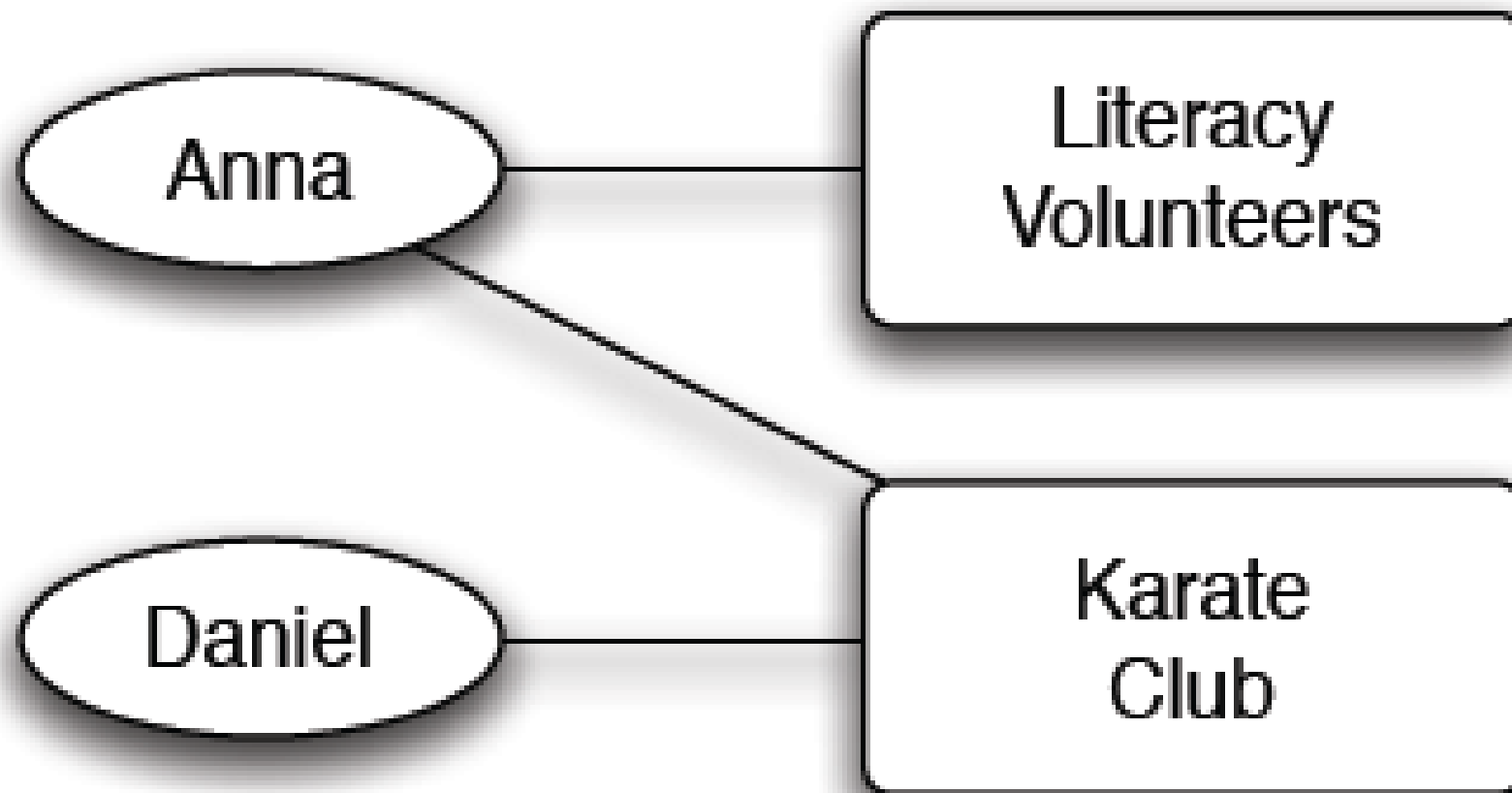
- 端点颜色相同的边如果太多，则同质性迹象明显
- 多少算“太多”？（什么是基准？—“混合得比较均匀”）
  - 考虑不同颜色节点的占比，随机给节点着色的情况

# 逆同质性社会网络的一个例子

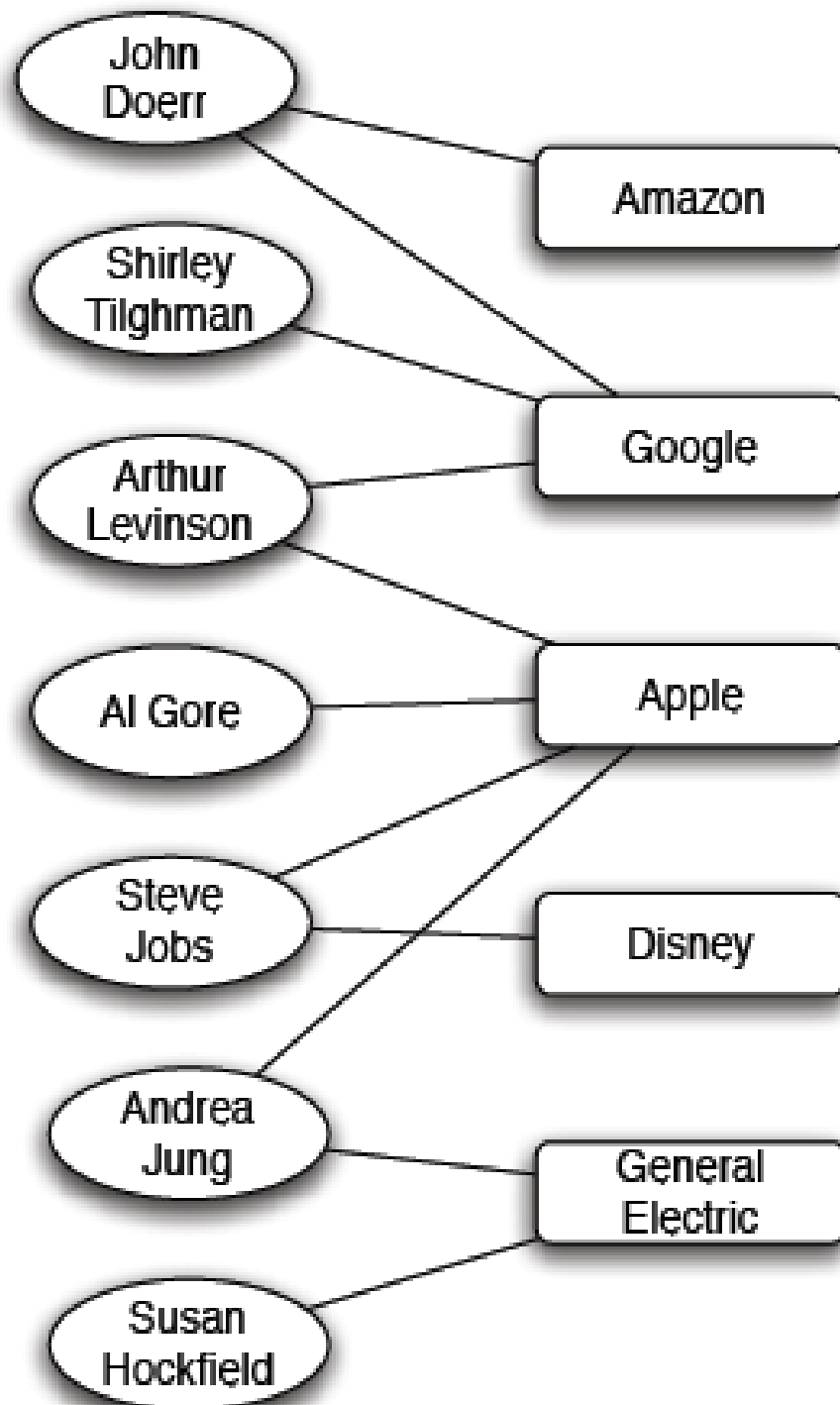




# 归属网络：描述人与相遇机会的关系



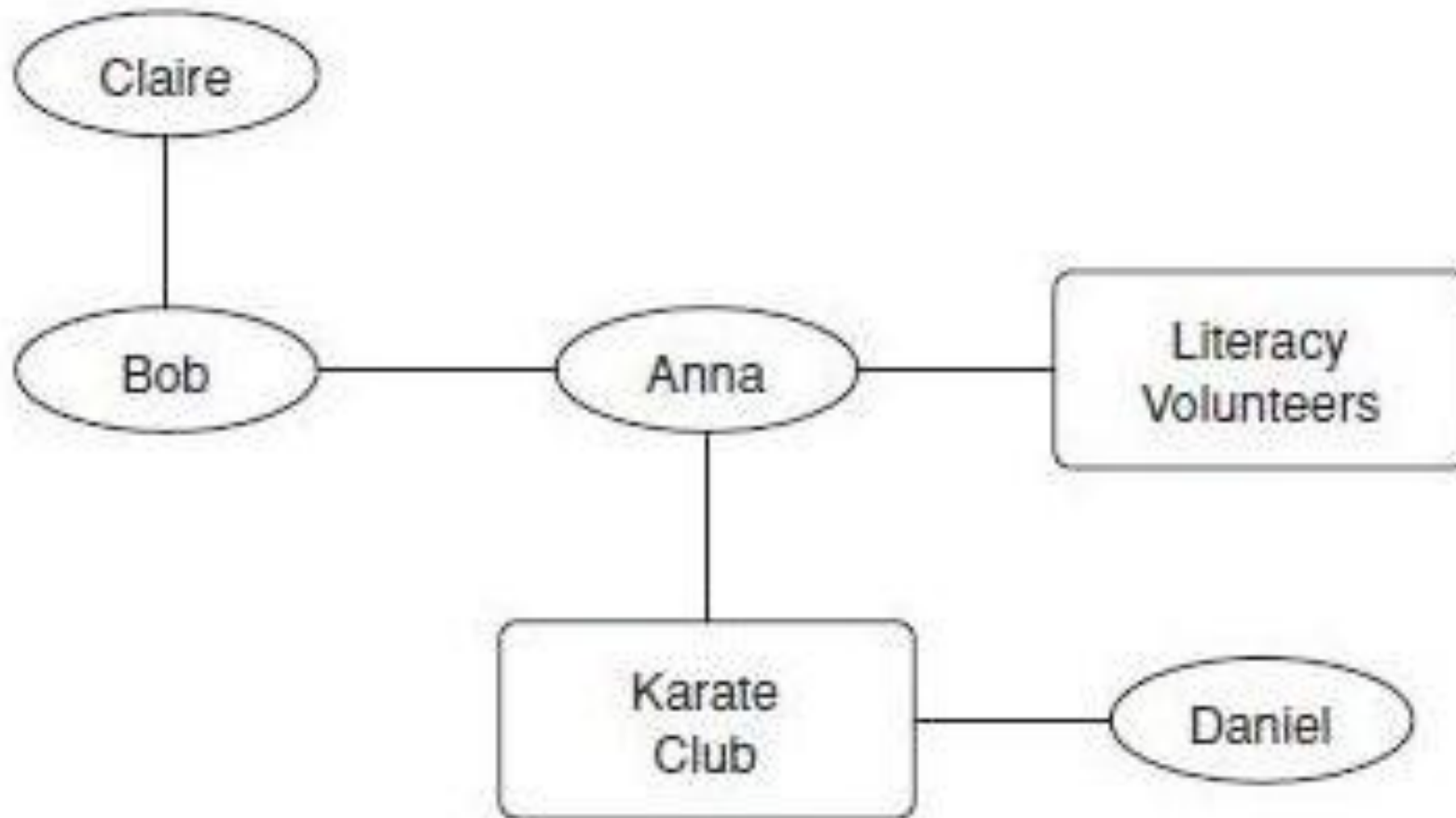
# 归属网络的一个实际例子



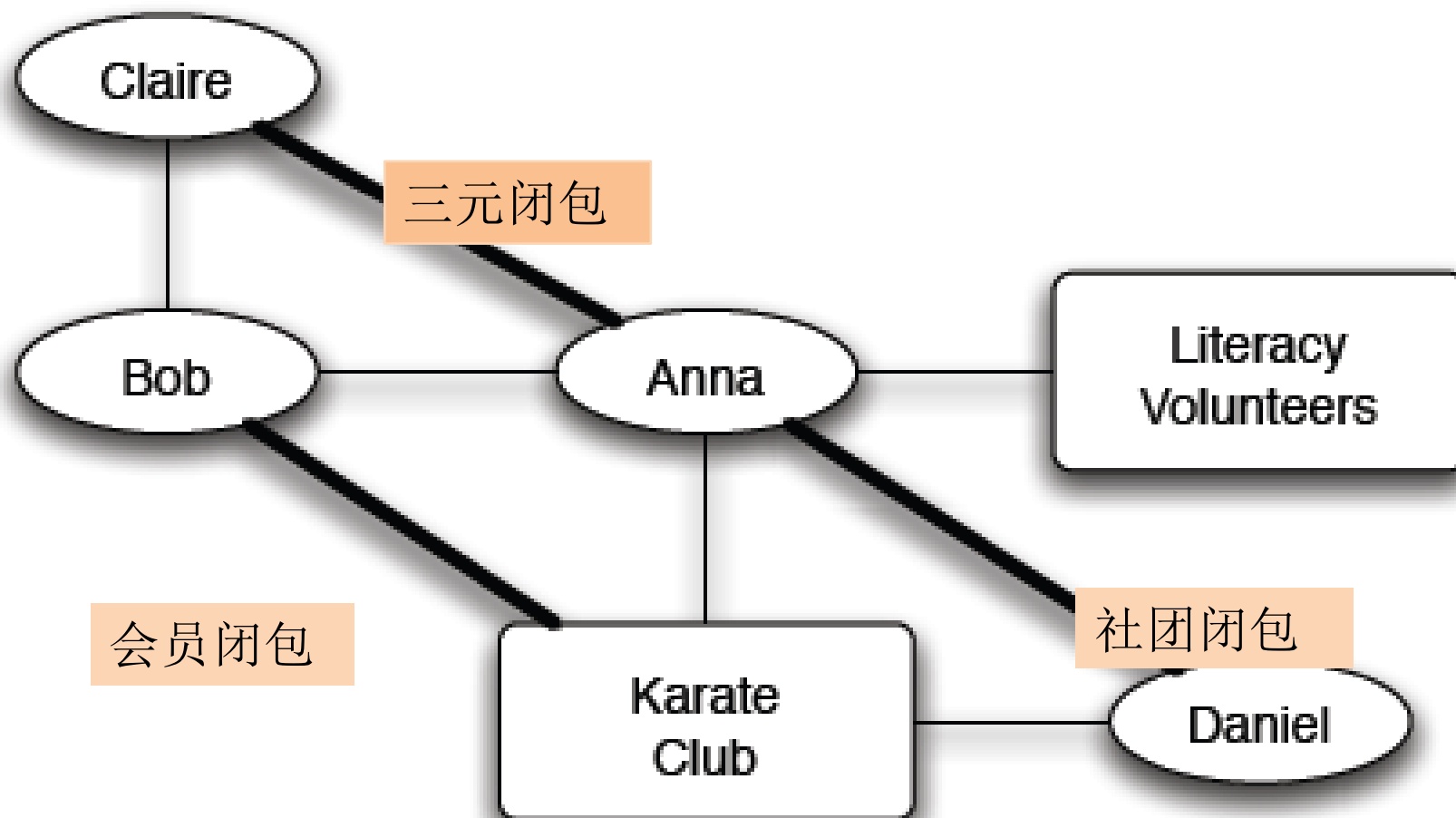
- 名人可能在多家公司的董事会中兼职，分析这种结构，以及他们之间的个人关系，对理解公司的行为有帮助
- 这种现象也可能造成一些丑闻



# 社会归属网：考察同质性的一个工具

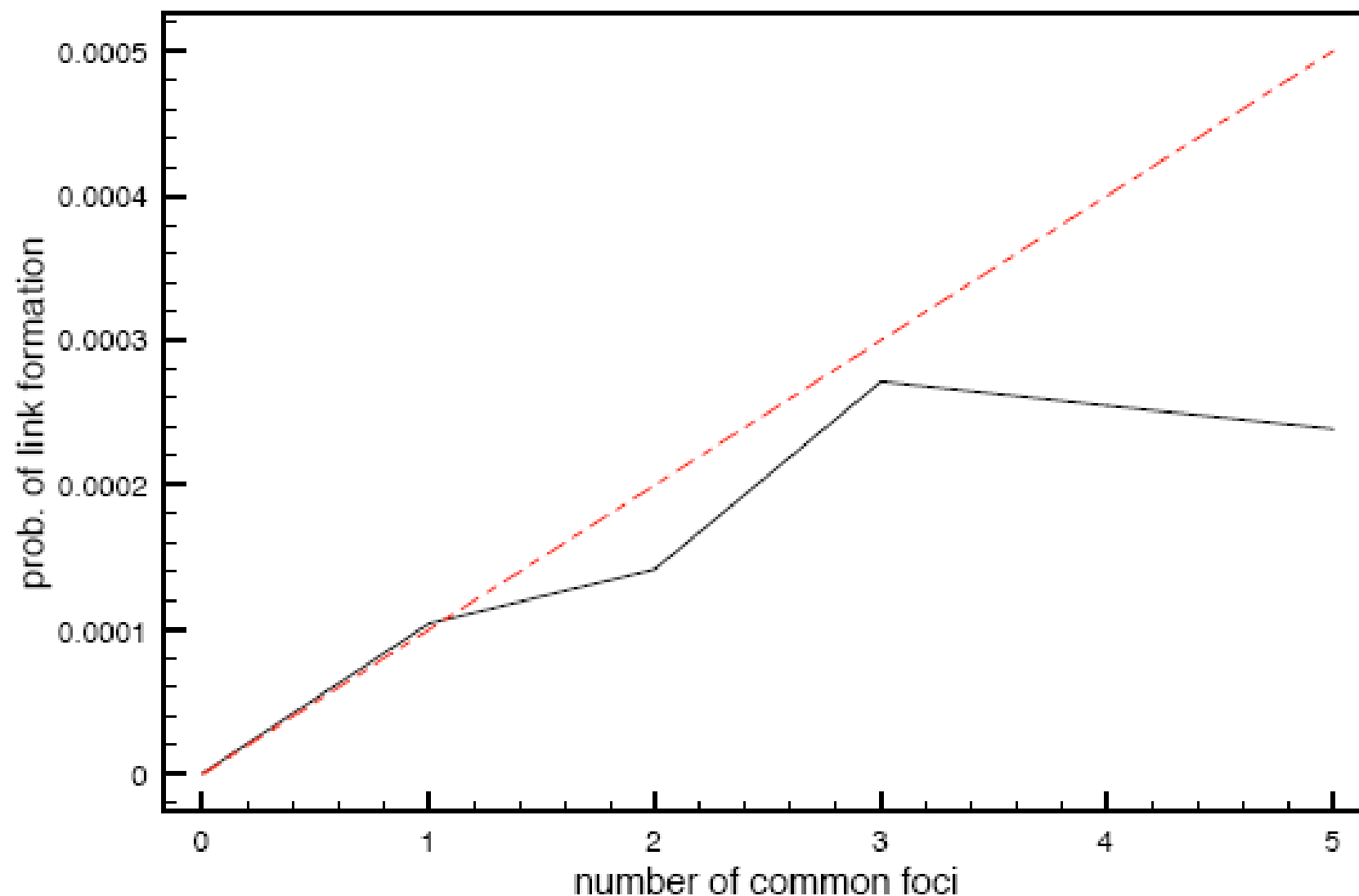


# 社会归属网： 三元闭包、社团闭包、会员闭包可同时显现



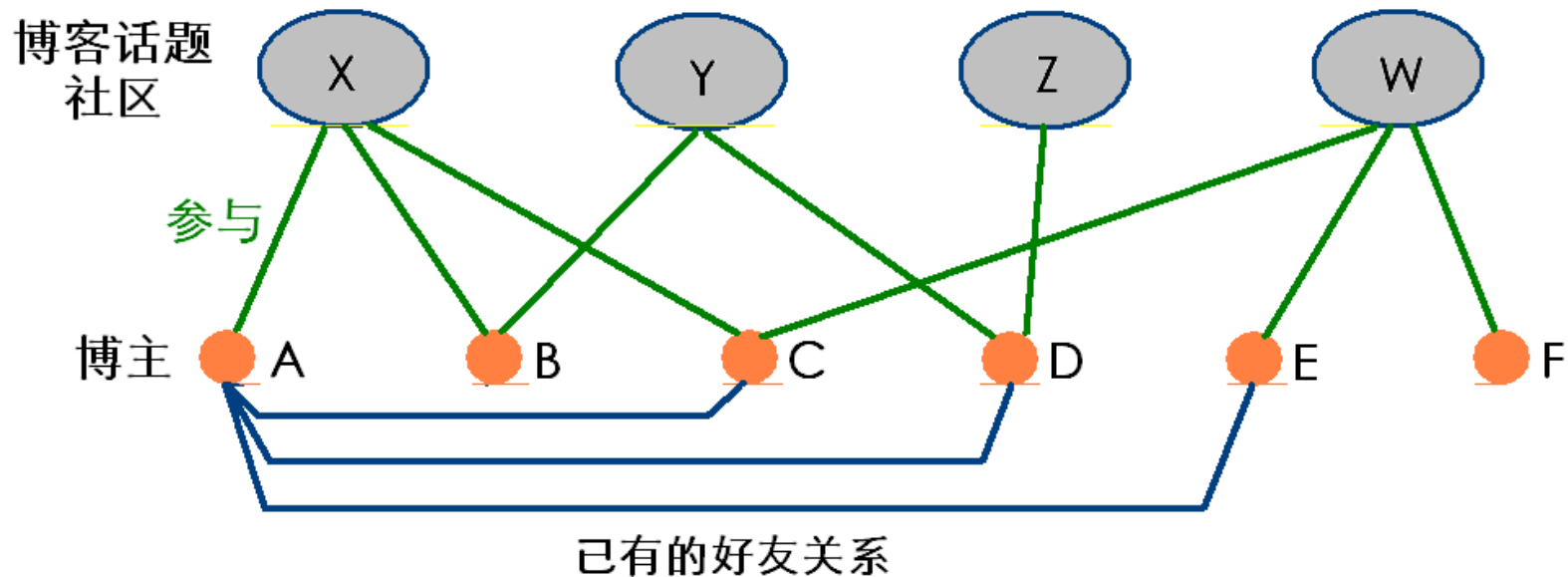
- 社团闭包→选择；会员闭包→社会影响

# 社团闭包验证：基于共同兴趣建立联系的概率



# 利用在线数据支持同质性研究

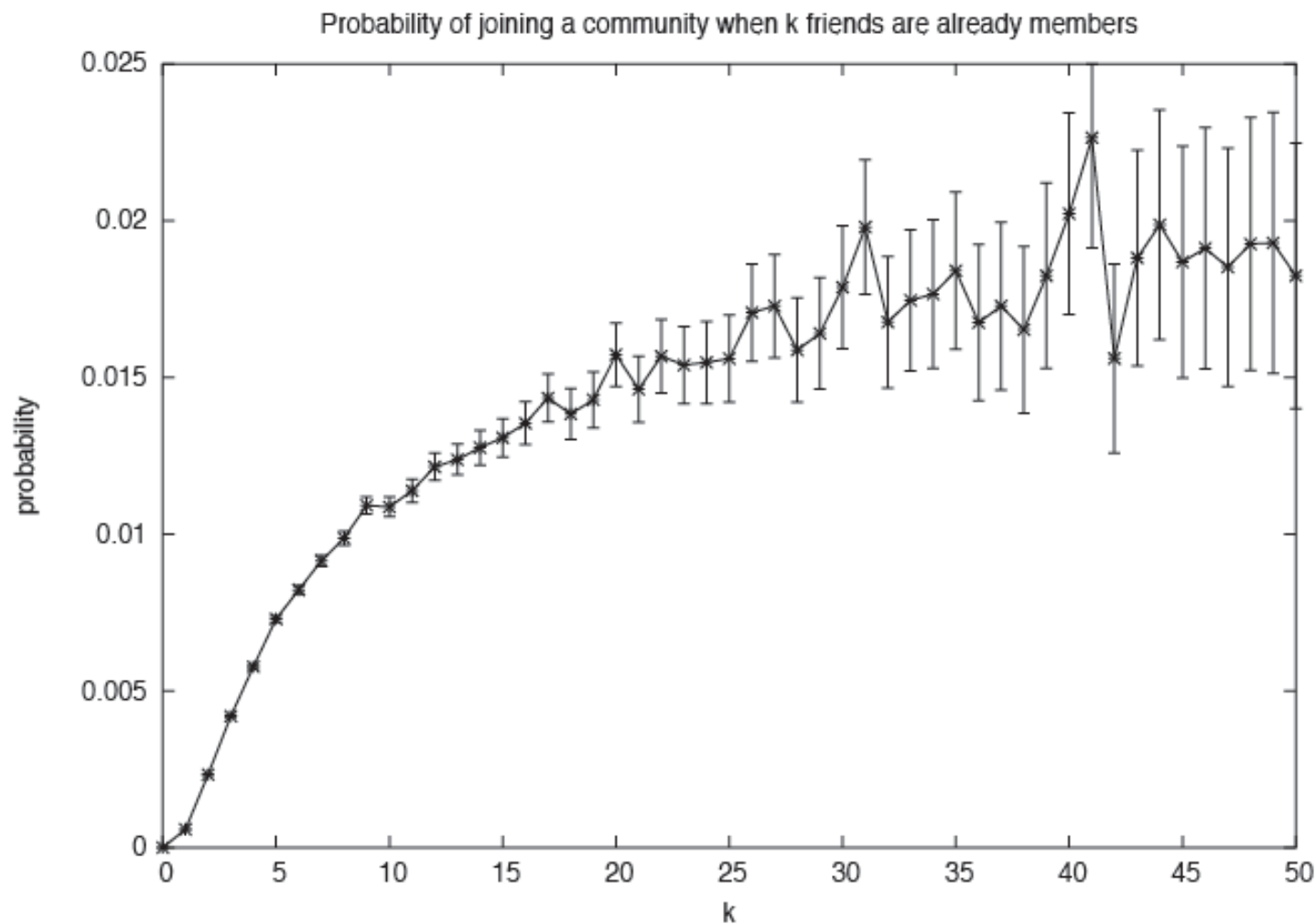
- 定义（社会归属网）：人、社团；人和人的关系，人和社团的关系
  - 社团：某些人共同参与的活动



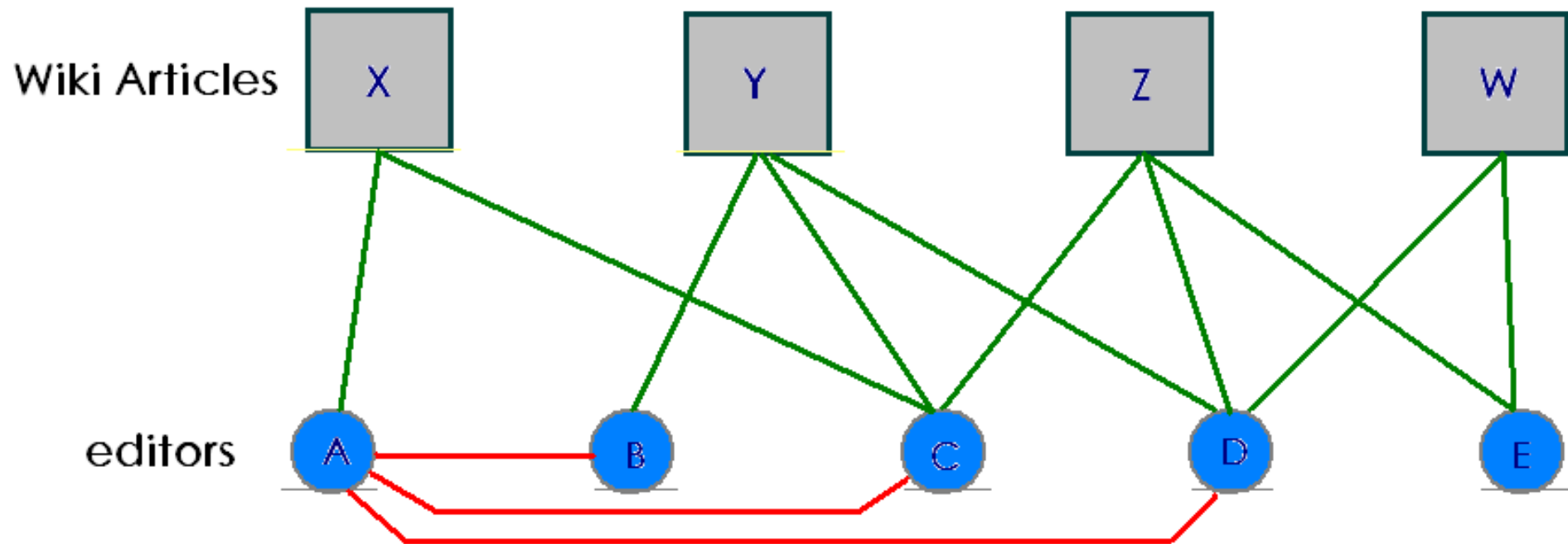
A加入Y? Z? W? 的可能性?

D D C,E

# 会员闭包（社会影响）： 基于相关朋友数，参与一个博客话题社区的概率

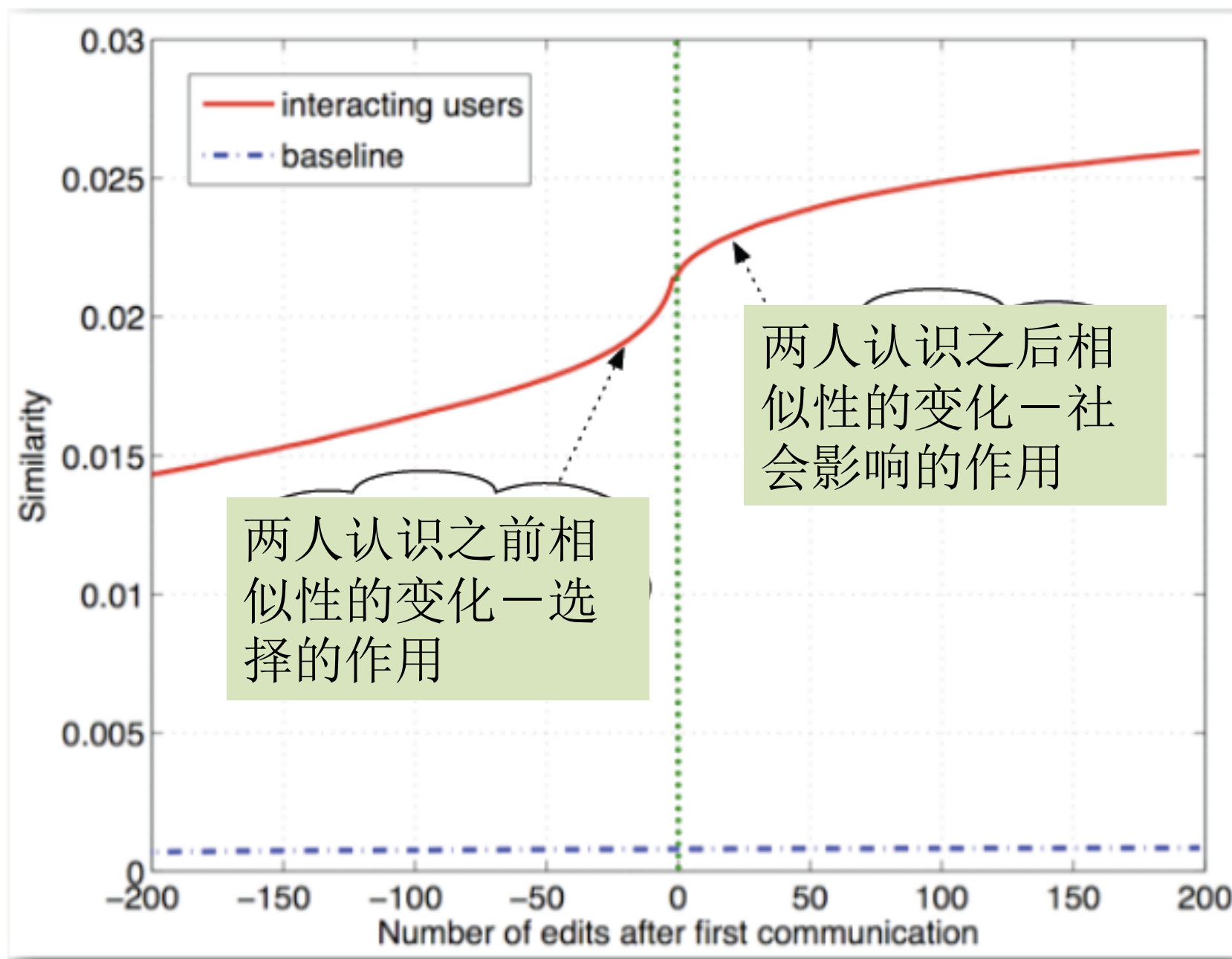


# 利用在线数据研究选择与社会影响的互动



- 两个编辑之间相似性的变化与“选择”和“社会影响”的关系
  - 定义“相似性” — 编辑文章（与 / 或）
- 没有联系（通信）之前，相似主要因为选择；达到足够相似度时则容易发生联系，然后社会影响开始对相似性提高起作用

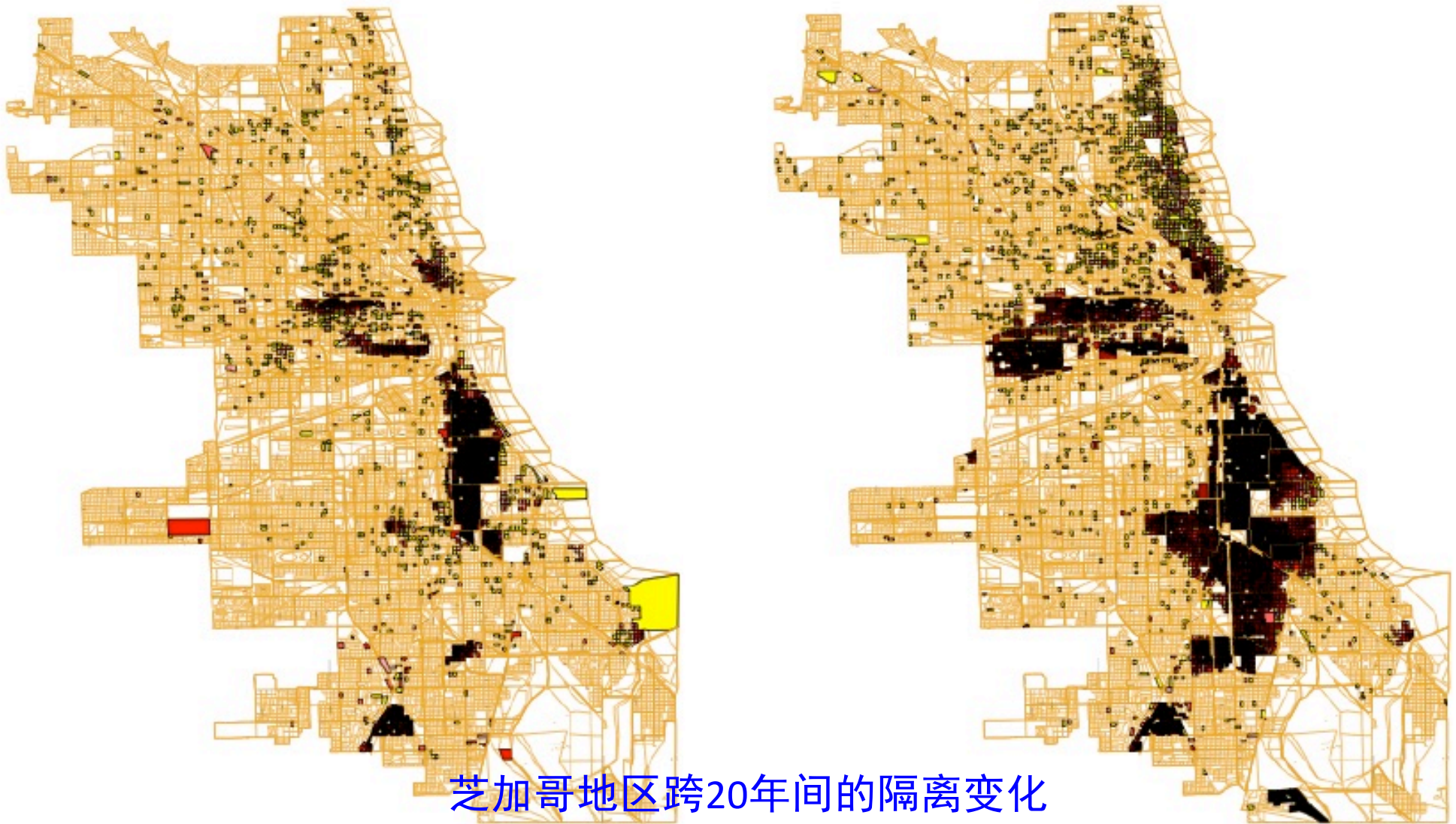
# 相似性、选择与社会影响





# 隔离 (segregation)

- 同质性影响下发生的过程

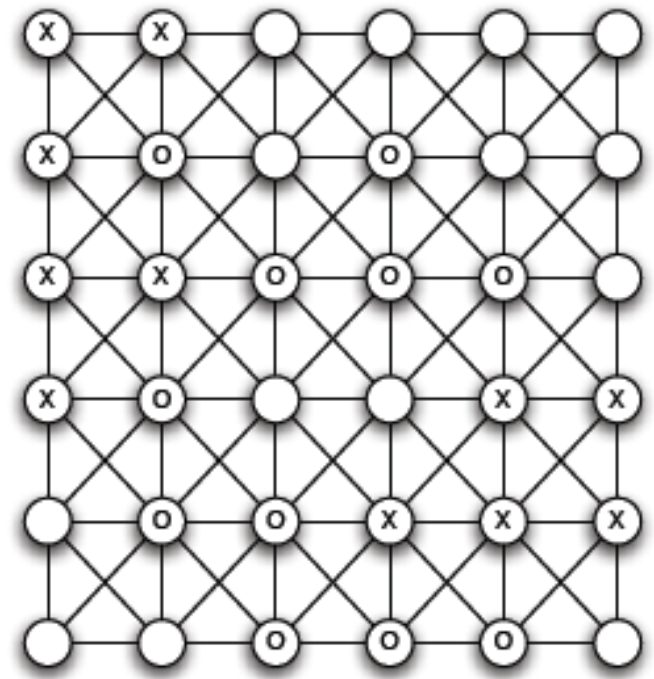


芝加哥地区跨20年间的隔离变化

# 谢林模型示意

- 解释隔离现象（宏观，全局）并不一定是个人刻意选择（微观，局部）的结果

X	X				
X	O		O		
X	X	O	O	O	
X	O			X	X
	O	O	X	X	X
		O	O	O	

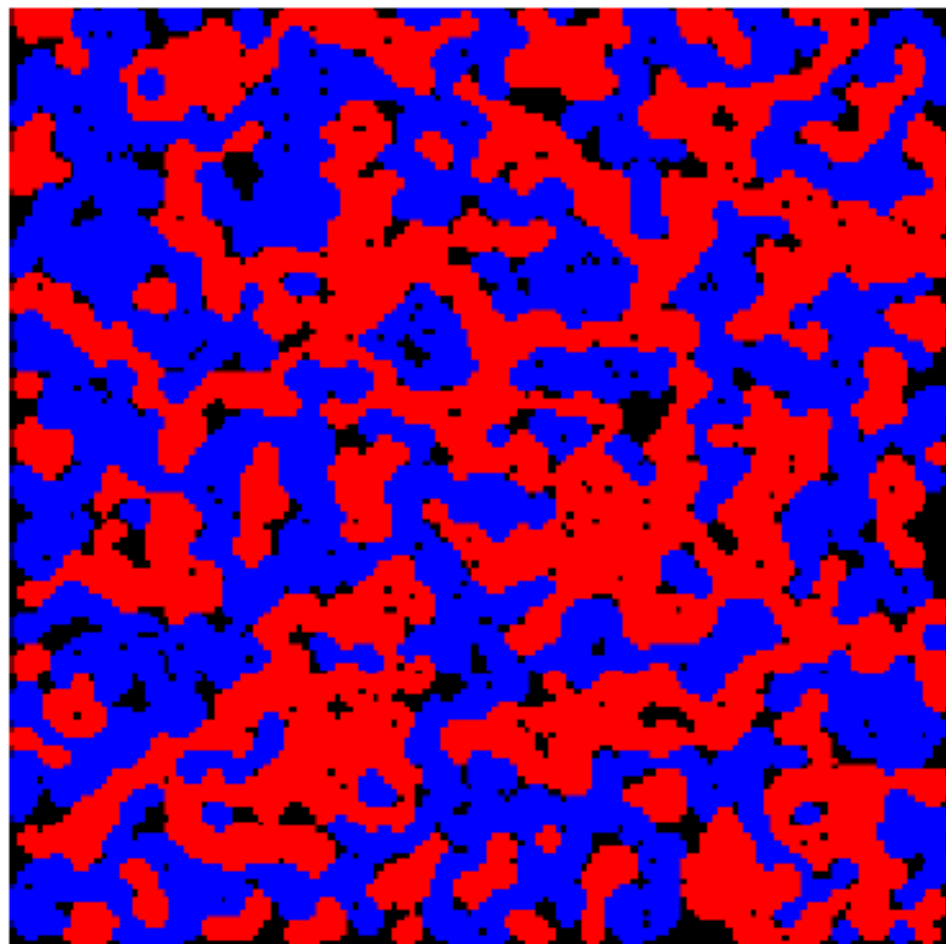
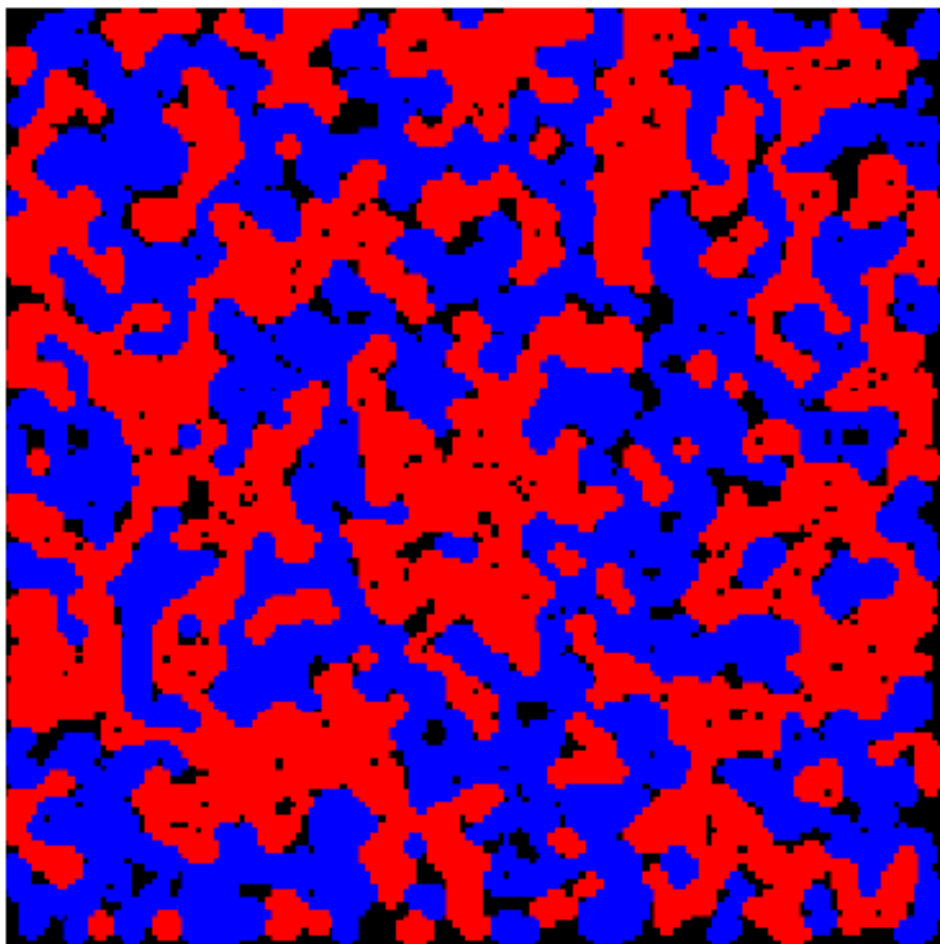


# 谢林模型 示意

X1*	X2*				
X3	O1*		O2		
X4	X5	O3	O4	O5*	
X6*	O6			X7	X8
	O7	O8	X9*	X10	X11
		O9	O10	O11*	

X3	X6	O1	O2		
X4	X5	O3	O4		
	O6	X2	X1	X7	X8
O11	O7	O8	X9	X10	X11
	O5	O9	O10*		

# 谢林模型的运行（两次）

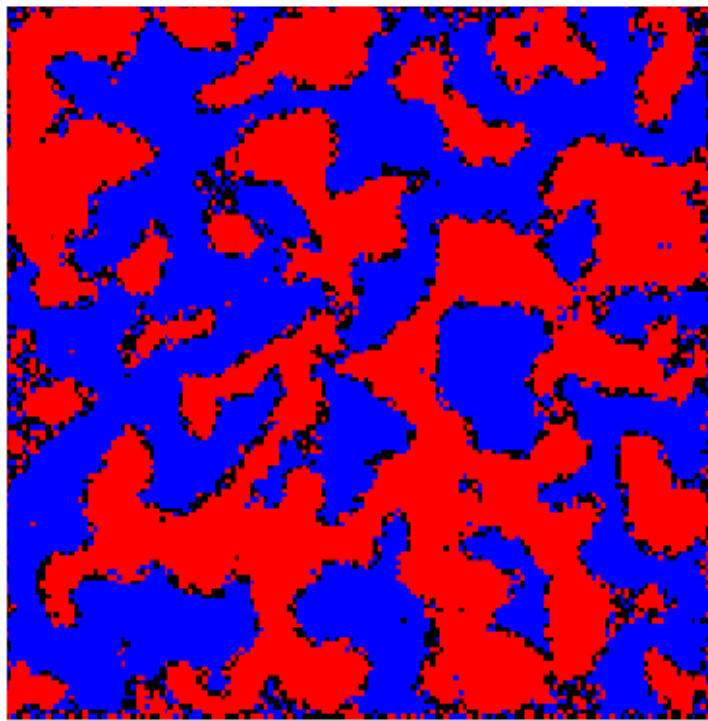


- 150行 \* 150列，各10000， $t=3$ ；随机初始化

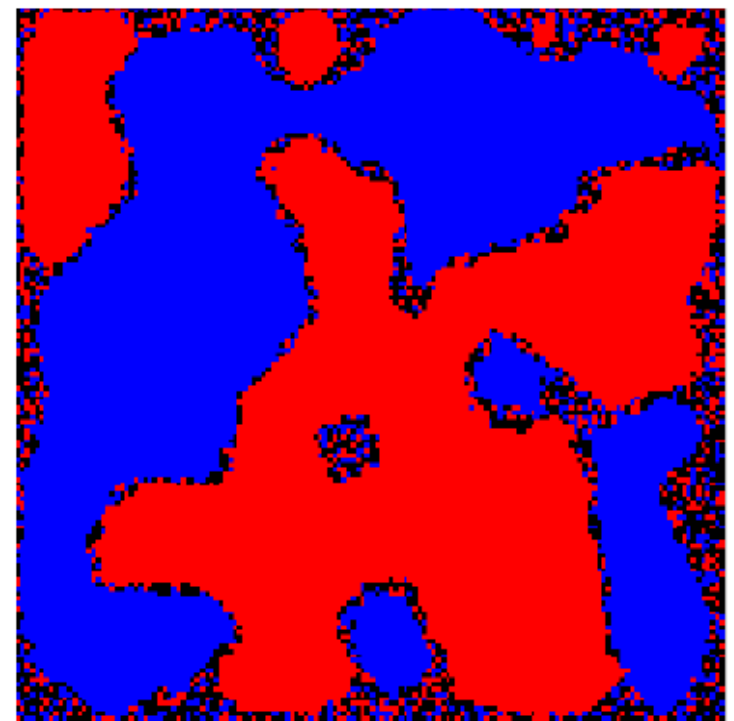
# 局部要求的一种无隔离的满足

X	X	O	O	X	X
X	X	O	O	X	X
O	O	X	X	O	O
O	O	X	X	O	O
X	X	O	O	X	X
X	X	O	O	X	X

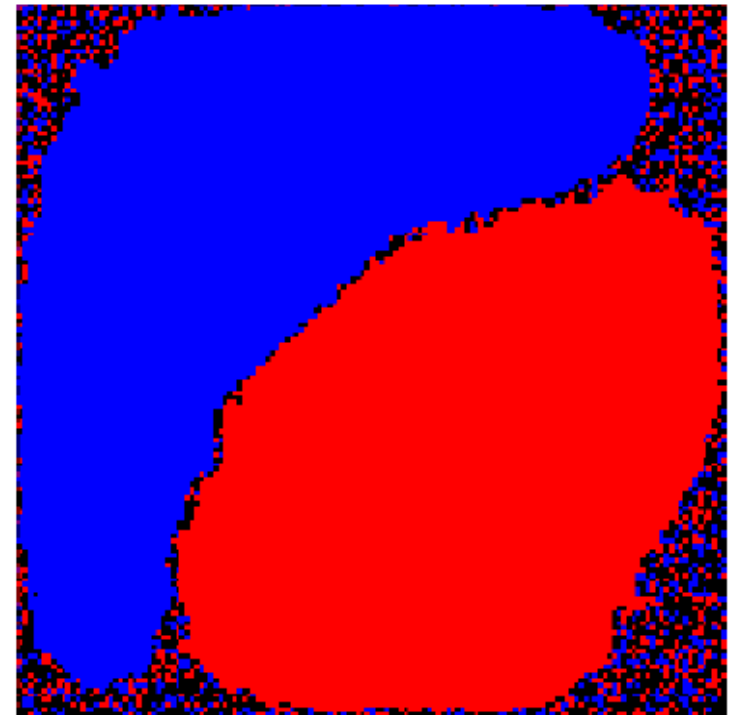
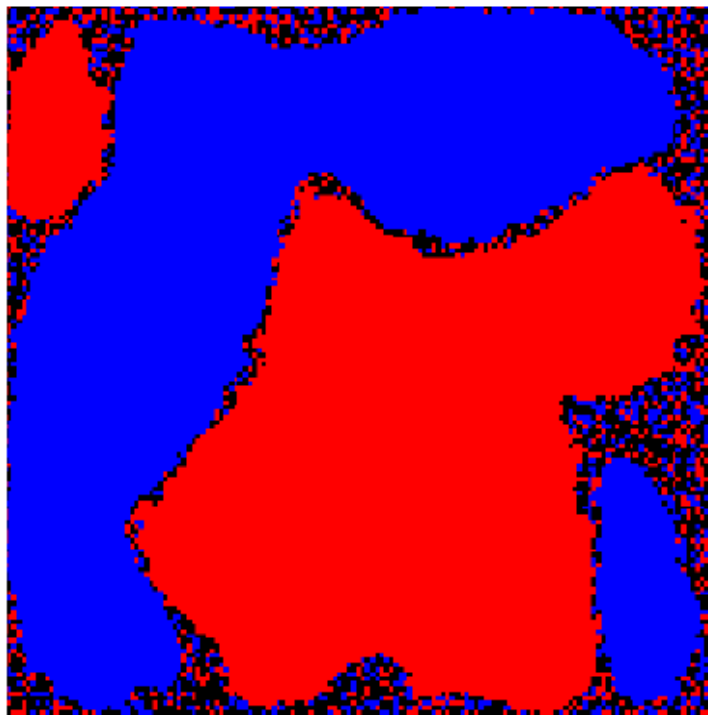
$$T=4$$



(a) *After 20 steps*



(b) *After 150 steps*



# 小节

- 同质性概念及其对社会网络的影响
- 社会网络中同质性迹象的评估
- 基于社会归属网的概念，结合OSN数据分析，考察同质性现象的方法
- 在同质性研究中的计算机模拟方法（以隔离为例）



# 作业

- 第4章 3,4
- Or
- 实现谢林模型并提交源代码及运行结果