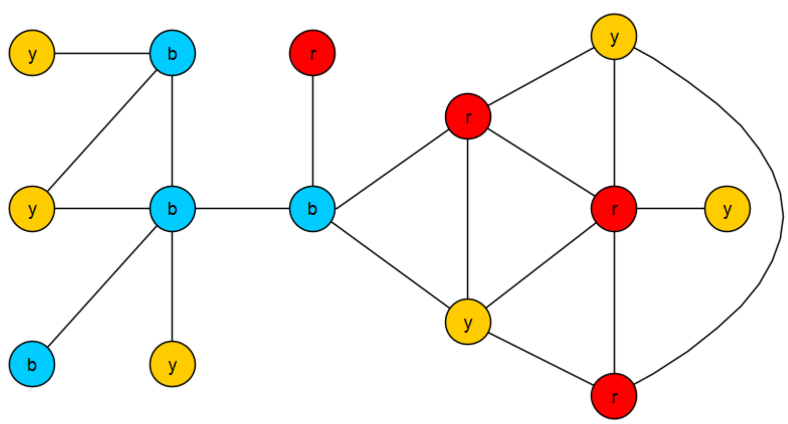
Prof. Dr. Markus Strohmaier
Dr. Fabian Flöck
Ivan Smirnov, Ph.D.
Tobias Schumacher

Pen & Paper Exercise 6

Social Networks

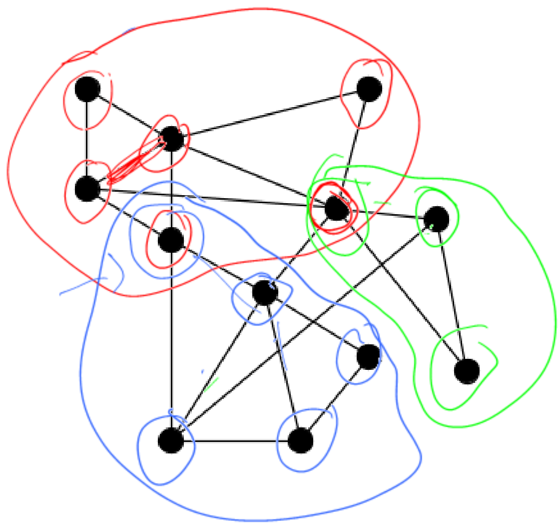
1 Label Propagation

Assume that after a few iterations of the label propagation algorithm, we are given the graph that is presented in the following diagram, in which only the labels *yellow* (y), *red* (r) and *blue* (b) are left:



Compute the next two iteration loops of the label propagation algorithm, assuming the nodes are always iterated over in top-down, then left-to-right order (i.e. first the three left-most nodes, with the top one (y) first, middle one (y) second, bottom one (b) third, then the blue node in top row, etc). What can you say regarding the eventual final outcome of the label propagation?

2 Clique Percolation



Perform the clique percoplation algorithm for $k = 3$ on the graph above. I.e., circle all the communities that you can find. Are there any overlapping communities?

Prof. Dr. Markus Strohmaier
Dr. Fabian Flöck
Ivan Smirnov, Ph.D.
Tobias Schumacher

3 Homophily

In a survey of heterosexual couples living in San Francisco, the fractions of couples with each combination of ethnic groups was estimated to be as follows:

		Women				
		Black	Hispanic	White	Other	Total
Men	Black	~~0.26~~	~~0.02~~	~~0.03~~	~~0.01~~	0.32
	Hispanic	0.01	0.16	0.06	0.02	0.25
	White	0.01	0.02	0.31	0.03	0.37
	Other	0.01	0.01	0.02	0.02	0.06
	Total	0.29	0.21	0.42	0.08	

Assuming these results to be representative of the network of relationships, compute the modularity of the network with respect to ethnicity. What do you conclude about homophily in this network?

$$M = \sum_c \left(\frac{L_c}{L} - \left(\frac{k_c}{2L}\right)^2 \right)$$

↳ $e_c = \frac{L_c}{L}$ ≙ observed fraction of links within community c

↳ $a_c = \frac{k_c}{2L}$ ≙ fraction of ends of edges that attach to a node in community c.

$$M = \sum_c (e_c - a_c^2)$$

Compute these values for all communities $c \in \{B, W, H, O\}$:

$e_B = 0.26$ $\quad a_B = 0.26 + 0.5 \cdot (4 \cdot 0.01 + 0.02 + 0.03) = 0.305$

ends of edges

$e_W = 0.31$ $\quad a_W = 0.31 + 0.5 \cdot (0.01 + 2 \cdot 0.02 + 2 \cdot 0.03 + 0.06) = 0.23$

$e_H = 0.16$ $\quad a_H = 0.16 + 0.5 \cdot (2 \cdot 0.01 + 3 \cdot 0.02 + 0.06) = 0.395$

$e_O = 0.02$ $\quad a_O = 0.02 + (3 \cdot 0.01 + 2 \cdot 0.02 + 0.03) = 0.07$

↳ $M = \sum_c (e_c - q_c^2) = \ldots = 0.44375$ → this is a high modularity ↳ strong homophily

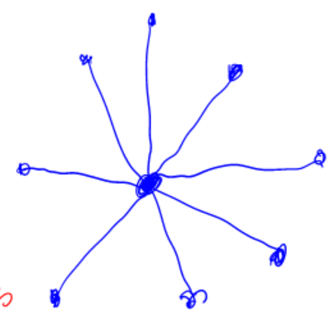
Prof. Dr. Markus Strohmaier
Dr. Fabian Flöck
Ivan Smirnov, Ph.D.
Tobias Schumacher

4 Degree Assortativity

Consider a star graph, i.e. a graph with N nodes, which contains one central node that is connected to all $N-1$ other nodes, which contrarily have degree one.

a) For this network, compute
   (i) the degree distribution p_k,
   (ii) the probability q_k that when moving along a randomly chosen link we find at its end a node with degree k,
   (iii) the values of the degree correlation function $k_{nn}(k)$ for all degrees k that occur in this network,
   (iv) the degree assortativity coefficient r.

b) Is this network assortative or disassortative? Justify your answer!



a)
$$p_k = \begin{cases} \frac{1}{N} & k = N-1, \\ 1-\frac{1}{N} & k = 1, \\ 0 & \text{else} \end{cases}$$

→ only two degrees occurring in the network, $\underbrace{k=1}_{\text{central node}}$ and $\underbrace{k=N-1}_{\text{all } N-1 \text{ outer node}}$

b)
$$q_k = \begin{cases} \frac{1}{2} & k = N-1, \\ \frac{1}{2} & k = 1, \\ 0 & \text{else} \end{cases}$$

→ all edges have one end with degree $k=N-1$, and one end with degree $k=1$

c) $k_{nn}(k) = \sum_{k'} k' \cdot P(k' \mid k)$

↳ in this graph $P(k' = N-1 \mid k=1) = 1 = P(k'=1 \mid k = N-1)$

↳ $k_{nn}(1) = N-1$, $k_{nn}(N-1) = 1$

d) $r = \dfrac{S_1 S_e - S_2^2}{S_1 S_3 - S_2^2}$ where

$S_e = 2\sum_{(i,j)\in E} k_i k_j = 2 \cdot \underbrace{(N-1)}_{\#\text{edges}} \cdot \underbrace{1}_{\text{outer node}} \cdot \underbrace{(N-1)}_{\text{central node}} = 2 \cdot (N-1)^2$

$S_1 = \sum_i k_i = (N-1) \cdot 1 + 1 \cdot N-1 = 2N-2$

$S_2 = \sum_i k_i^2 = (N-1) \cdot 1^2 + 1 \cdot (N-1)^2 = N(N-1)$

$S_3 = \sum_i k_i^3 = (N-1) \cdot 1^3 + 1 \cdot (N-1)^3 = (N-1)(N^2-2N+2)$

plug in and compute ↓ $= -1$ → graph is strongly dis-assortative