

Pagerank Lab Report

by Marcus Kindberg

Transition probabilities

The transition matrix for the graph described in three.txt is

$$P = \begin{pmatrix} \frac{3}{80} & \frac{29}{48} & \frac{77}{240} & \frac{3}{80} \\ \frac{3}{80} & \frac{3}{80} & \frac{71}{80} & \frac{3}{80} \\ \frac{77}{240} & \frac{77}{240} & \frac{3}{80} & \frac{77}{240} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix} = \begin{pmatrix} 0.0375 & 0.6042 & 0.3208 & 0.0375 \\ 0.0375 & 0.0375 & 0.8875 & 0.0375 \\ 0.3208 & 0.3208 & 0.0375 & 0.3208 \\ 0.25 & 0.25 & 0.25 & 0.25 \end{pmatrix}$$

and its 10th power is

$$P^{10} = \begin{pmatrix} 0.0375 & 0.6042 & 0.3208 & 0.0375 \\ 0.0375 & 0.0375 & 0.8875 & 0.0375 \\ 0.3208 & 0.3208 & 0.0375 & 0.3208 \\ 0.25 & 0.25 & 0.25 & 0.25 \end{pmatrix}$$

The transition matrix P can be broken down into $P = \alpha(H + D) + \frac{1-\alpha}{n}\mathbf{1}$, where $H = \begin{pmatrix} 0 & \frac{2}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 1 & 0 \\ \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 0 & 0 & 0 \end{pmatrix}$ and $D = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$.

Results

The following table gives the top hits, i.e., the 5 first vertices of each graph sorted by page rank, using $\alpha = \frac{85}{100}$.

three.txt	2 (36.4%)	1 (28.0%)	0 (17.9%)	3 (17.9%)
tiny.txt	0 (27.0%)	1 (26.0%)	3 (24.6.9%)	2 (15.0%)
medium.txt	6 (5.9%)	22 (5.7%)	9 (4.0%)	13 (3.9%)
wikipedia.txt	1 (40.0%)	2 (35.6%)	4 (6.5%)	5 (3.4%)
p2p-Gnutellao8-mod.txt	367 (0.24%)	249 (0.135%)	145 (0.131%)	264 (0.017%)

The following table gives the number of random walk steps and (scalar) multiplications needed for each graph until the results were stable to within 2 decimal places.

Graph	# transitions	# multiplications
three.txt	83	8
tiny.txt	115	16
medium.txt	113	12
wikipedia.txt	90	31
p2p-Gnutellao8-mod.txt	170	4