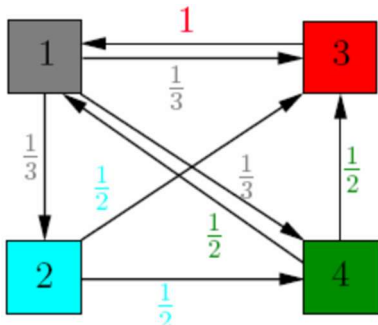


The goal of a search engine is to rank results in such a way that the top results are more relevant to the user. In this submission I'll talk describe PageRank, the algorithm used by Google to rank their search results, and which works by solving an eigenvalue problem.

The algorithm models the pages A, B, C, ... as nodes of a directed graph, and an edge A->B corresponds to an hyperlink in A leading to B. Then, assume that the person surfing the web clicks mindlessly, hopping to any hyperlink in the page with equal probability.



We can organize this as an adjacency matrix:

$$A = \begin{bmatrix} 0 & 0 & 1 & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 0 & 0 \end{bmatrix}$$

The PageRank algorithm works by solving the eigenvalue problem $Av = v$ (for instance, $v = [12 \ 4 \ 9 \ 6]$ in the example). The length of the components of the eigenvector v represent the probability that the person ends up in that component, so you can use that to rank search results (so rank $1 > 3 > 4 > 2$).