

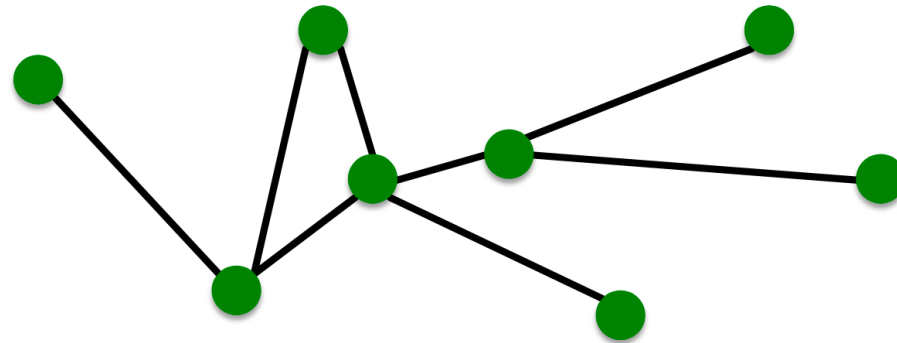
PageRank

Hongchang Gao

Spring 2024

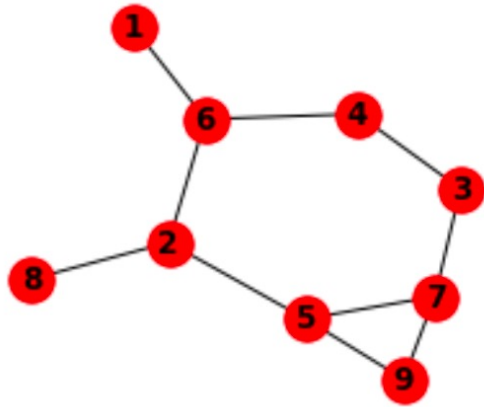
Graph Terminology

- Components of a Graph
 - Nodes/vertices
 - Edges/links
 - Graph/Network

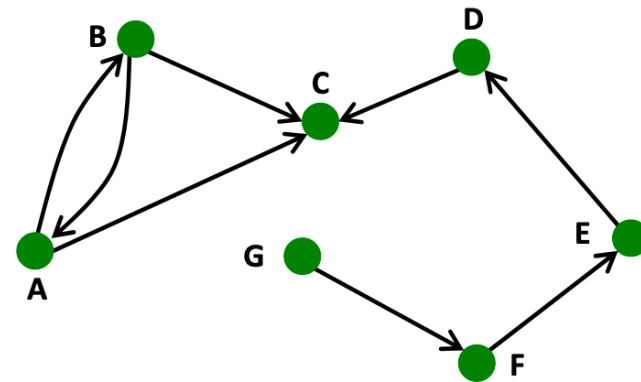


Graph Terminology

- Types of Graphs
 - Undirected Graph: links are undirected
 - Friendship on Facebook
 - Directed Graph: links are directed
 - Following on Twitter



Undirected graph

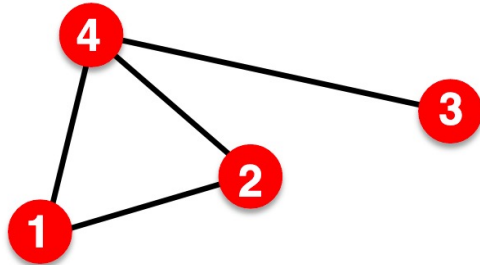


Directed graph

Graph Terminology

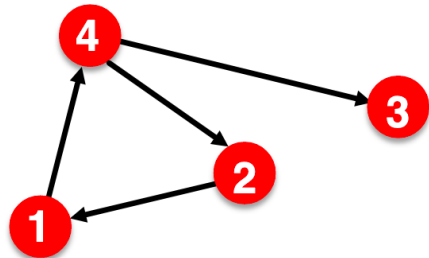
- Adjacency matrix

Undirected



$$A = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$
$$A_{ij} = A_{ji}$$
$$A_{ii} = 0$$

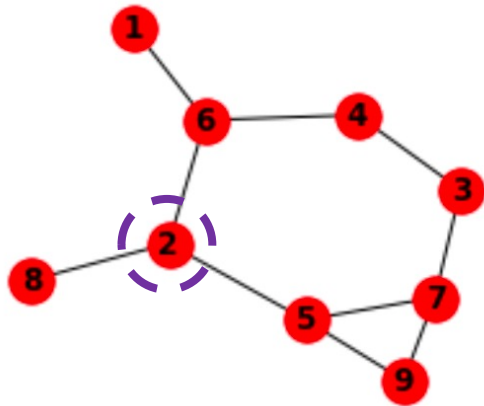
Directed



$$A = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$
$$A_{ij} \neq A_{ji}$$
$$A_{ii} = 0$$

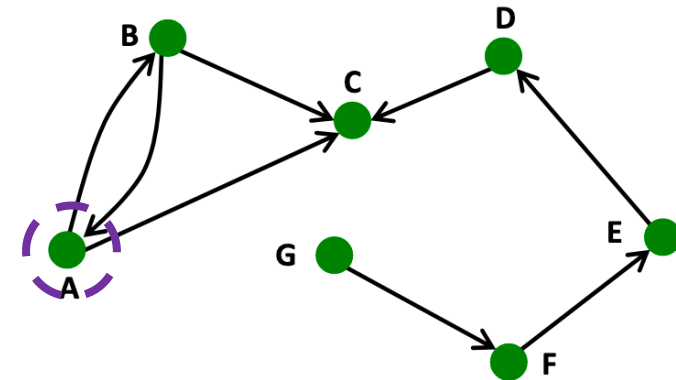
Graph Terminology

- **Node degrees** of undirected graph
 - The number of edges adjacent to a node



Node 2: $d = 3$

- **Node degrees** of directed graph
 - In-degree: the number of **head ends** adjacent to a node
 - Out-degree: the number of **tail ends** adjacent to a node

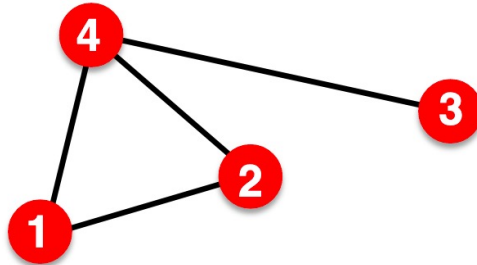


Node A: $d_{in}=1, d_{out}=2$

Graph Terminology

- Node degrees

Undirected



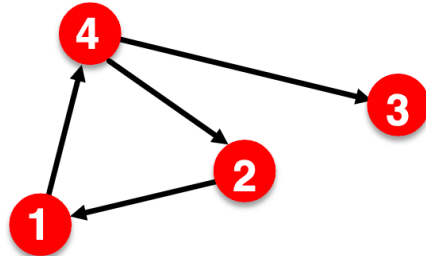
$$A = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

$$A_{ij} = A_{ji}$$

$$A_{ii} = 0$$

Node 2: $1+0+0+1=2$

Directed



$$A = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

$$A_{ij} \neq A_{ji}$$

$$A_{ii} = 0$$

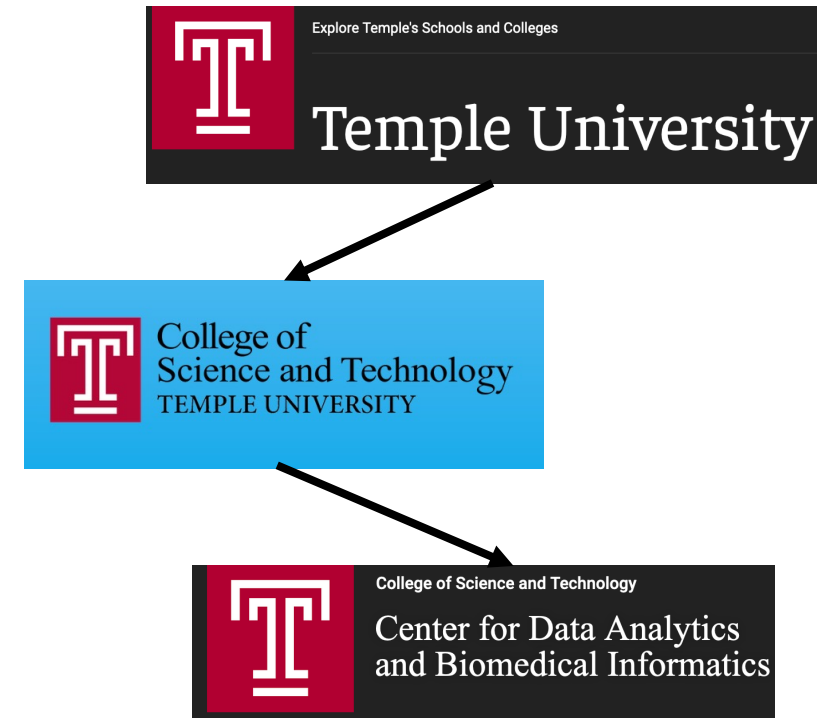
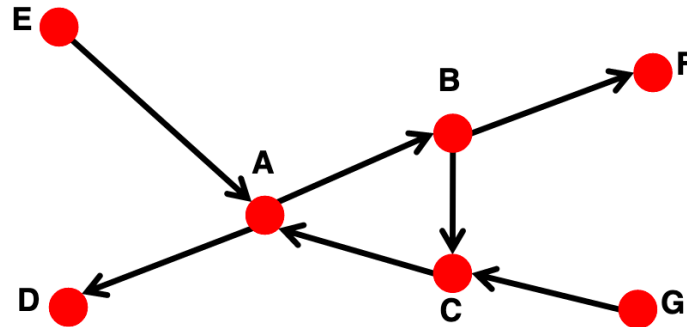
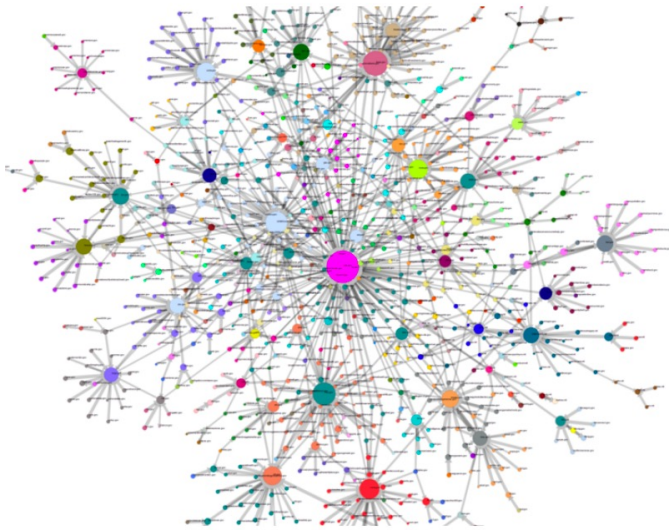
Node 4:

Indegree (column sum): $1+0+0+0=1$

Outdegree (row sum): $0+1+1+0=2$

PageRank

- The web is a graph
 - Nodes: web pages
 - Edges: hyperlinks



PageRank

- All web pages are not equally “important”
 - Some webpages should be assigned more priority than others, for being more important
 - Which node (webpage) is important?



🔍 vaccine ✕

Google Search

I'm Feeling Lucky

scdhec.gov › covid19 › covid-19-vaccine-appointments ⋮

COVID-19 Vaccine Appointments | SCDHEC

Make an appointment with a **vaccine** provider to guarantee you'll receive your shot. If you're receiving the Pfizer-BioNTech or Moderna **vaccines**, you need to make ...

[COVID-19 Vaccine Provider](#) · [Understanding the Vaccination...](#) · [Herd Immunity](#)

www.cdc.gov › vaccines ⋮

Vaccines and Immunizations | CDC

Vaccine Communications. Past Announcements · Newsletters · Immunization MMWRs. email_03Get Email Updates. To receive email updates about this page, ...

[COVID-19 Vaccination](#) · [Immunization Schedules](#) · [The Basics](#) · [Adults](#)

www.cdc.gov › coronavirus › 2019-ncov › vaccines ⋮

When You've Been Fully Vaccinated | CDC

6 days ago — Recommendations on what activities people can do after they have been fully **vaccinated**, including how to gather safely with **vaccinated** and ...

vaccinefinder.org ⋮

VaccineFinder - Find COVID-19 vaccine locations near you

VaccineFinder helps you find clinics, pharmacies, and other locations that offer COVID-19 **vaccines** in the United States. In some states, information may be ...

Pa

• Ra

• Pa

• Se

Multi Search

university

Search

Next! [national parks]

10 results

clustering on

Search

Query: university

11 Results Returned

Showing Results From 0 to 10

Stanford University Homepage

74.79%

http://www.stanford.edu/

4k - 3/29/93 - 01/03/97

Stanford University: Portfolio Collection

65.78%

http://www.stanford.edu/home/administration/portfolio.html

3k - 3/29/93 - 01/03/97

University of Illinois at Urbana-Champaign

73.26%

http://www.uiuc.edu/

13k - 12/30/95 - 01/03/97

Indiana University

68.38%

http://www.indiana.edu/

1k - 09/23/95 - 01/05/97

University of California, Irvine

68.07%

http://www.uci.edu/

3k - 12/30/95 - 01/03/97

University of Minnesota

67.05%

http://www.umn.edu/

0k - 12/18/95 - 01/03/97

Iowa State University Homepage

66.66%

http://www.iastate.edu/

3k - 12/18/95 - 01/03/97

The University of Michigan

66.35%

http://www.umich.edu/

1k - 3/29/93 - 01/03/97

Mississippi State University

66.35%

http://www.msstate.edu/

3k - 3/29/93 - 01/03/97

Northwestern University: NUIInfo

66.15%

http://www.nwu.edu/

3k - 12/14/95 - 01/05/97

next 10

Optical Physics at the University of Oregon

Oregon Center for Optics in Science and Technology. Department of Physics, University of Oregon, Eugene OR 97403. Research Groups: Carmichael Group....
<http://optics.uoregon.edu/> - size 1K - 16 Dec 96

Carnegie Mellon University - Campus Networking

Departments. Data Communications. Data Communications is responsible for installing and maintaining all on campus networking equipment and all of...
<http://www.net.cmu.edu/> - size 4K - 19 Aug 95

Wesleyan University Computer Science Group Home Page

Computer Science Group. Wesleyan University. Welcome to the home page of the Computer Science Group at Wesleyan University. We are administratively within.
<http://www.cs.wesleyan.edu/> - size 3K - 15 Apr 96

Keio University Shonan Fujisawa Campus (SFC)

B3\$N%Z!EFnF#Bt%-9c%9s%Q%9 (B(SFC) \$B\$N (BWWW \$B% \$BCm0U=q\$- (B \$B\$rFI\$s\$G\$!/\$@5\$ \$!# (B. Nihongo | English. SFC \$B>pJs (B. [\$B%a%G%#%"%9s%?!*...
<http://www.sfc.keio.ac.jp/> - size 3K - 5 Feb 97

School of Chemistry, University of Sydney

The School of Chemistry. School of Chemistry, University of Sydney, NSW 2006 Australia International Phone: +61-2-9351-4504 Fax: +61-2-9351-3329 Australia.
<http://www.chem.su.oz.au/> - size 4K - 25 Feb 97

Mankato State University

The Campus Athletics, Campus Tour, Bookstore, Maps, Current Events... Admission & Registration Admissions, Financial Aid, Registrar's, Graduate...
<http://www.mankato.msus.edu/> - size 3K - 27 Nov 96

St. Ambrose University

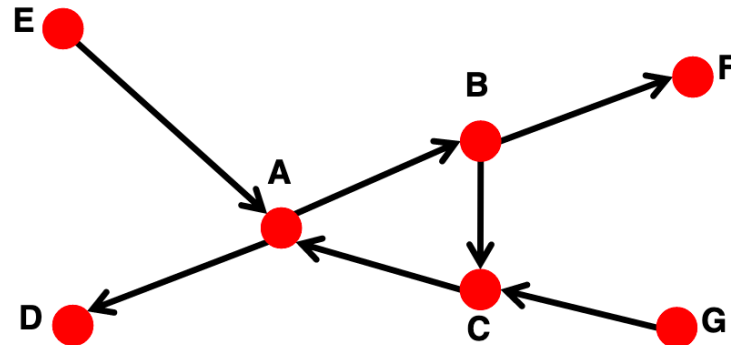
Main Index: Academic Departments. Administrative Services. Campus News. Computing Services. Galvin Fine Arts Center. Internet Connections. Library...
<http://www.sau.edu/> - size 3K - 4 Feb 97

University of Washington ECSEL Projects

ery

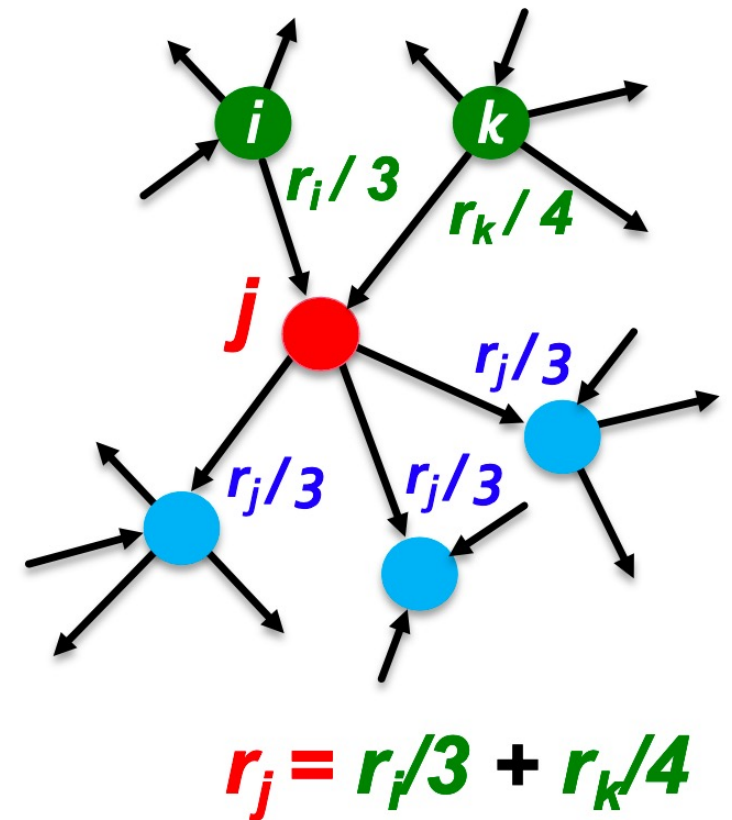
PageRank

- Compute the importance of nodes in a graph
 - Idea: Links as votes
 - Page is more important if it has more links
- Use in-links as votes
 - How to use votes to compute the importance score???
 - Q: E and C may be different. C may be more important.
 - How to differentiate their importance when they vote for A?



PageRank

- A vote from an important page is worth more:
 - Each link's vote is proportional to the **importance** of its source page
 - If page i with importance r_i has d_i out-links, each link gets r_i / d_i votes
 - Page j 's own importance r_j is the sum of the votes on its in-links



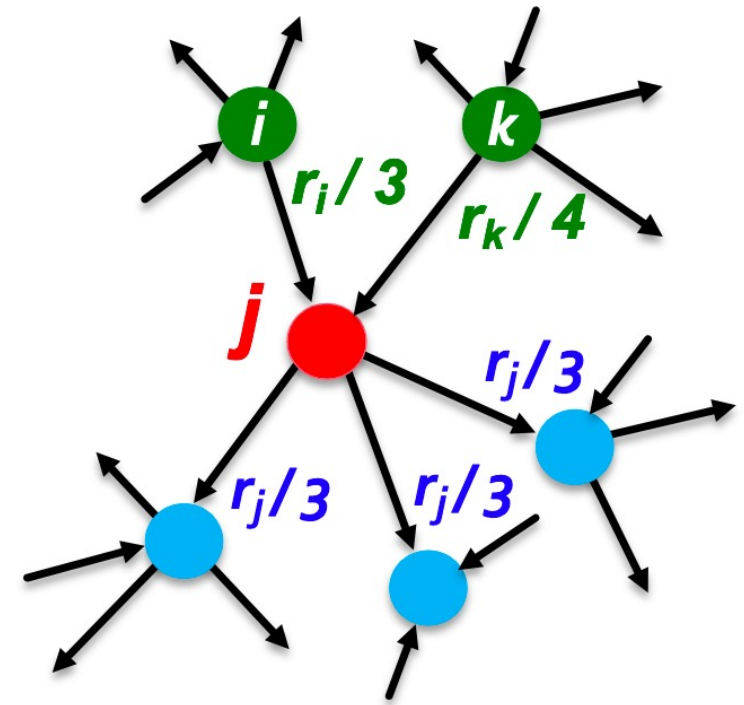
PageRank

- Formally, the importance score of each node is

$$r_j = \sum_{i \rightarrow j} \frac{r_i}{d_i}$$

d_i ... out-degree of node i

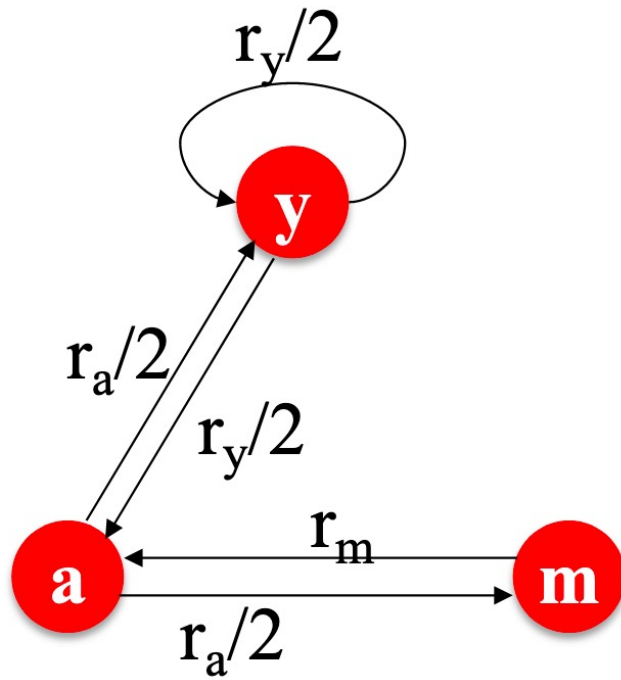
A page is important if it is pointed by other important pages



$$r_j = r_i/3 + r_k/4$$

PageRank

- Example



$$r_j = \sum_{i \rightarrow j} \frac{r_i}{d_i}$$

d_i ... out-degree of node i

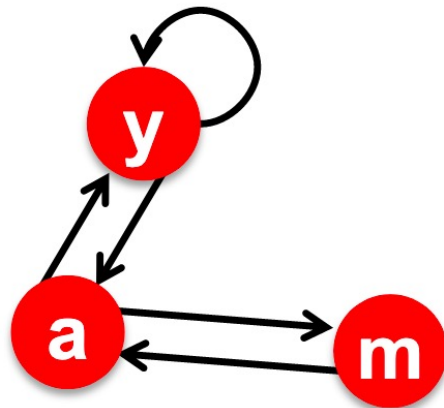
$$r_y = r_y/2 + r_a/2$$

$$r_a = r_y/2 + r_m$$

$$r_m = r_a/2$$

PageRank

- Matrix Form



Graph

Out-degree 2, 2, 1

	y	a	m
y	1	1	0
a	1	0	1
m	0	1	0

Adjacency matrix

	r_y	r_a	r_m
r_y	$\frac{1}{2}$	$\frac{1}{2}$	0
r_a	$\frac{1}{2}$	0	1
r_m	0	$\frac{1}{2}$	0

Transition matrix M

$$r_j = \sum_{i \rightarrow j} \frac{r_i}{d_i}$$

d_i ... out-degree of node i

$$r_y = r_y/2 + r_a/2$$

$$r_a = r_y/2 + r_m$$

$$r_m = r_a/2$$

$$\begin{bmatrix} r_y \\ r_a \\ r_m \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 1 \\ 0 & \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} r_y \\ r_a \\ r_m \end{bmatrix}$$

$\mathbf{r} \quad \mathbf{M} \quad \mathbf{r}$

PageRank

- Property of the transition matrix M
 - Column sum is 1
- Property of the rank vector r
 - r_i is the importance score of page i

$$\sum_i r_i = 1$$

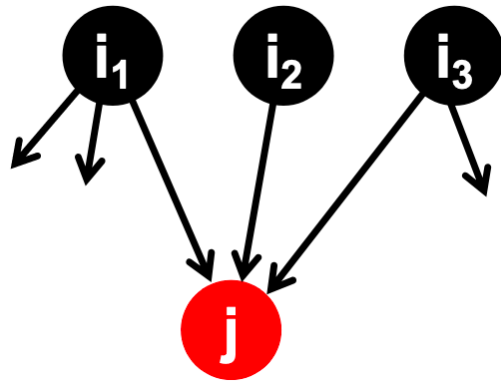
$$\begin{array}{|c|} \hline r_y \\ \hline r_a \\ \hline r_m \\ \hline \end{array} = \begin{array}{|ccc|} \hline \frac{1}{2} & \frac{1}{2} & 0 \\ \hline \frac{1}{2} & 0 & 1 \\ \hline 0 & \frac{1}{2} & 0 \\ \hline \end{array} \begin{array}{|c|} \hline r_y \\ \hline r_a \\ \hline r_m \\ \hline \end{array}$$

$\mathbf{r} \qquad \mathbf{M} \qquad \mathbf{r}$

PageRank

- Interpretation

- At time t , the user is on page i
- At time $t+1$, the user follows an out-link from i uniformly at random
- Ends up on some page j linked from i



$$r_j = \sum_{i \rightarrow j} \frac{r_i}{d_{\text{out}}(i)}$$

r_y
r_a
r_m

$$=$$

$\frac{1}{2}$	$\frac{1}{2}$	0
$\frac{1}{2}$	0	1
0	$\frac{1}{2}$	0

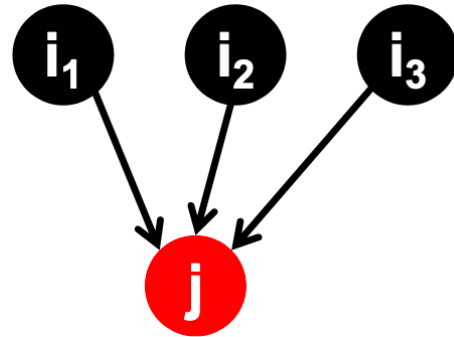
r_y
r_a
r_m

\mathbf{r} \mathbf{M} \mathbf{r}

PageRank

- Interpretation (continue):
 - Define $\mathbf{p}(t)$ is a probability distribution over pages
 - At time $t+1$, we have

$$\mathbf{p}(t + 1) = \mathbf{M} \cdot \mathbf{p}(t) \quad \text{Random walk}$$

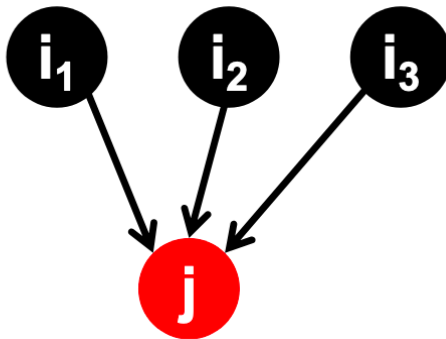


PageRank

- Solution of the importance score r :

$p(t)$ is **stationary distribution** of a random walk

$$p(t + 1) = M \cdot p(t)$$



Random walk



Stationary distribution

$$r = M \cdot r$$

r is the stationary distribution of the random walk

r is the eigenvector of the transition matrix M (with eigenvalue 1)

PageRank

- Solution of the importance score r :
 - Compute the eigenvector of the transition matrix M with eigenvalue 1
 - Use power iteration to compute the eigenvector efficiently

- Assign each node an initial page rank
- Repeat until convergence ($\sum_i |r_i^{t+1} - r_i^t| < \epsilon$)
 - Calculate the page rank of each node

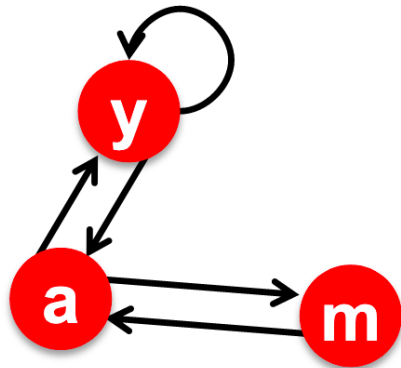


$$r_j^{(t+1)} = \sum_{i \rightarrow j} \frac{r_i^{(t)}}{d_i}$$

- Initialize: $\mathbf{r}^0 = [1/N, \dots, 1/N]^T$
- Iterate: $\mathbf{r}^{(t+1)} = \mathbf{M} \cdot \mathbf{r}^t$
- Stop when $|\mathbf{r}^{(t+1)} - \mathbf{r}^t|_1 < \epsilon$

PageRank

- Example



	y	a	m
y	$\frac{1}{2}$	$\frac{1}{2}$	0
a	$\frac{1}{2}$	0	1
m	0	$\frac{1}{2}$	0

$$\begin{matrix} r_y \\ r_a \\ r_m \end{matrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 1 \\ 0 & \frac{1}{2} & 0 \end{bmatrix} \begin{matrix} r_y \\ r_a \\ r_m \end{matrix}$$

r ***M*** ***r***

$$\begin{bmatrix} r_y \\ r_a \\ r_m \end{bmatrix} = \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix} \begin{bmatrix} 1/3 \\ 3/6 \\ 1/6 \end{bmatrix} \begin{bmatrix} 5/12 \\ 1/3 \\ 3/12 \end{bmatrix} \begin{bmatrix} 9/24 \\ 11/24 \\ 1/6 \end{bmatrix} \dots \begin{bmatrix} 6/15 \\ 6/15 \\ 3/15 \end{bmatrix}$$

Iteration 0, 1, 2, ...