

# PageRank

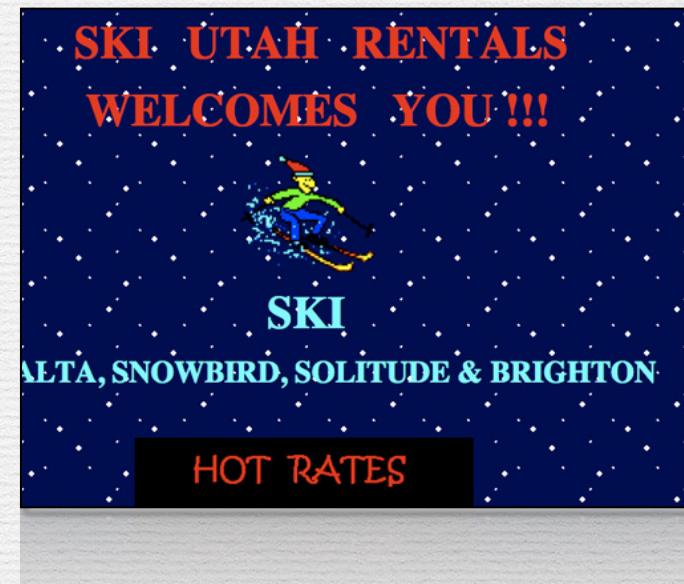
<https://www.youtube.com/watch?v=Quk88piD8PM>

[https://www.youtube.com/watch?v=LVV\\_93mBfSU](https://www.youtube.com/watch?v=LVV_93mBfSU)

# 1990's

Data

Webpages (URLs)



**Featured Areas**

Click the buttons below to view our featured sections!

**Hercules**

**McDonald's Japan**

**McWorld**  
Hey, IT COULD HAPPEN

Welcome to the McDonald's Official Web Site



Click on the [Golden Arches](#) to enter the main site!

Greetings boys and girls. You have arrived at the Zoo. A visual and literary landscape of [writings](#), music, and other things well suited to their neighborhood surroundings.

The Zoo is my exclusive outlet for my [writings](#). Want to awaken your soul? through the grinding super sounds of [Regurgitator](#), Brisbane's musical contribution to the world. All you lucky ladies and lads will all the best news from the most ear-pleasing including [Tea Party](#), [the Brown Hornet](#), [Muffcake](#) and [Fugazi](#).

"Random and Irrelevant Thoughts", and enter into the minds of me and my irrelevant words of wisdom in my [guestbook](#).

and give you what I couldn't find on the net. So help yourselves, enjoy and explore the manic world of my Zoo - [Zan](#).

Vote for this site in the Top10 Australian Music Web Sites

Your Name :   
Email Addr :

You must complete the Details above for your vote to count

**Smith and Goldsmith Inc.**

**SG**

Rental Properties  
PO Box 365  
Salem, Oregon  
97302-0655

**droom homes, lofts, duplexes and apartments.**

CELEBRATING OVER 50 YEARS of SERVICE  
Locally owned and owner operated.

Smith and Goldsmith has continually offered affordable, rental housing. With residential properties in Northeast and Southeast Salem areas, we focus on buildings well suited to their neighborhood surroundings. Places in which we'd like to (and have) lived! Smith and Goldsmith, we welcome the opportunity to show you the place you can call home!

Call Russell Smith, President  
Weekdays (503) 871-8432 9-5 Pacific time.  
Faxes are welcome and quickly replied.  
E-mail me at the link below.

[E-mail Me](mailto:smithandgoldsmith@yahoo.com)

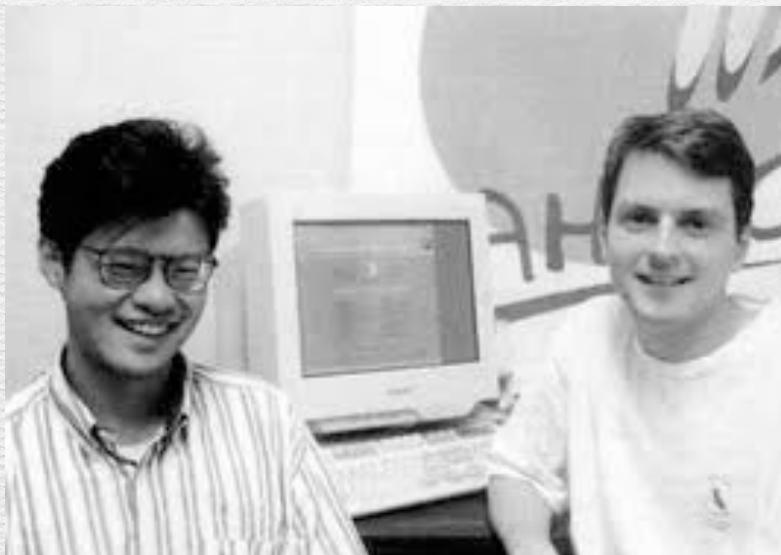


**Lofts at 388-400 Fairview Avenue SE offer**

# 1994

Data

Problem



Jerry Yang and David Filo

## Webpages (URLs)

Jerry and David's Guide  
to the World Wide Web

## Searching Webpage

### Yahoo

[Y Top](#) | [Up](#) | [Search](#) | [mail](#) | [Add](#) | [Help](#)

- [Art \(619\) NEW](#)
- [Business \(8546\) NEW](#)
- [Computers \(3266\) NEW](#)
- [Economy \(898\) NEW](#)
- [Education \(1839\) NEW](#)
- [Entertainment \(8814\) NEW](#)
- [Environment and Nature \(268\) NEW](#)
- [Events \(64\) NEW](#)
- [Government \(1226\) NEW](#)
- [Health \(548\) NEW](#)
- [Humanities \(226\) NEW](#)
- [Law \(221\) NEW](#)
- [News \(301\) NEW](#)
- [Politics \(184\) NEW](#)
- [Reference \(495\) NEW](#)
- [Regional Information \(4597\) NEW](#)
- [Science \(3289\) NEW](#)
- [Social Science \(115\) NEW](#)
- [Society and Culture \(933\) NEW](#)

*There are currently 31897 entries in the Yahoo database*

---

Some Other General Internet Directories:

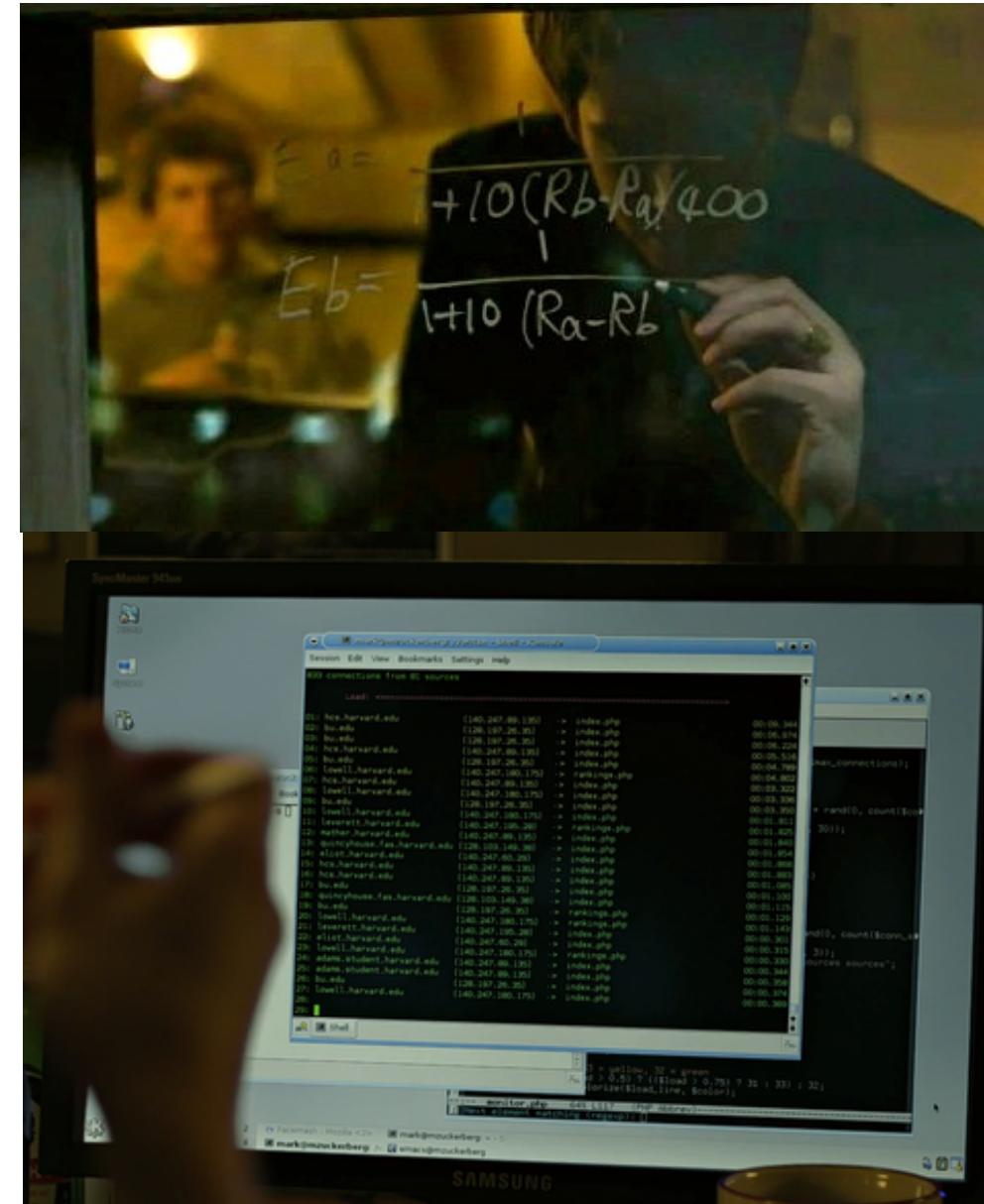
[ [WWW Virtual Library](#) \* [EINet Galaxy](#) \* [University of Michigan Clearinghouse](#) ]  
[ [GNN - Whole Internet Catalog](#) \* [Planet Earth](#) \* [Yanoff's Connections](#) ]

# Data Mining Example

Business Problem?

Math Problem?

Algorithm / Solution?



# Rating Problem (by Human or Machine)

**Trial Issue** INFER Click Here

investing in English antique furniture

Indian officials heighten security in Srinagar for Independence Day celebration

**CNN.com** Search CNN.com Find

CNN Sites myCNN Video | Audio Headline News Brief Free E-mail | Feedback

August 15, 2000 -- Updated 12:52 a.m. EDT, 0452 GMT

**BREAKING NEWS** Six cars of a Burlington Northern Santa Fe freight train have derailed in Montana, forcing evacuation of up to 400 people, a railway official reports. Three of the tanker cars contained a hazardous material, the official says. No leaks have been reported.

**CNN ON TV: Democratic Convention Coverage**

Watch Live on CNN (all times EDT)

- 7 a.m.: Bill Bradley's take on the convention
- 8 a.m.: Patrick Kennedy on his dad's address
- 3 p.m.: TalkBack Live on the women's vote

**DEMOCRATIC Convention Video** Watch interview with Hillary Clinton and other video highlights. Live coverage begins at 4:30 p.m. EDT

**In Other News:**

- Bad weather hampers Russian submarine rescue
- Tropical Storm Beryl bears down on Texas-Mexico border
- North and South Korea prepare for first family reunions since 1985
- China noncommittal on Arafat's timetable for Palestinian state

Clinton says Democrats hold keys to bright future

Saying the Democratic Party holds the keys to "progress and prosperity," President Clinton made the case for the election of his protege, Vice President Al Gore, in an emotional address to the party convention Monday night.

"Are we going to keep this progress, and

**myCNN** Headline News brief

**ZAN'S ZOO**

Greetings boys and girls. You have arrived at the Zoo. A visual and literary landscape of [writings](#), music, [interviews](#), and everything that will interest, excite, entertain and inform you all.

Want to broaden your mind? The Zoo is my exclusive outlet for my [writings](#). Want to awaken your soul? Allow me to take you through the grinding super sounds of [Regurgitator](#), Brisbane's musical contribution to the world. I'll also provide all you lucky ladies and lads with all the best news from the most ear-pleasing groups in the galaxy. Including [Tea Party](#), the [Brown Hornet](#), [Muffcake](#) and [Fugazi](#).

You can check out my "[Random and Irrelevant Thoughts](#)", and enter into the minds of me and my irrelevant buddies.

Feel free to [leave some words](#) of wisdom in my [guestbook](#).

I'm here to serve you and give you what I couldn't find on the net. So help yourselves, enjoy and open yourselves up to the manic world of my Zoo - Zan.

**Top10 Music Web Sites**

Vote for this site in the Top10 Australian Music Web Sites

Your Name :

Email Addr :

You must complete the Details above for your vote to count

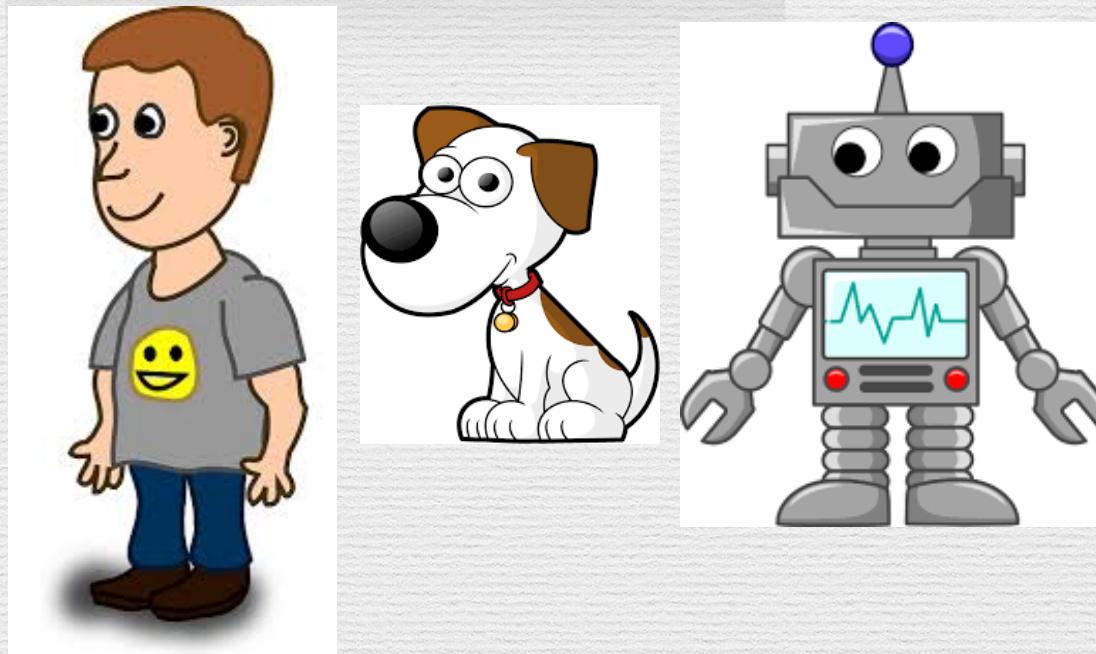
Voted Number One three weeks in a row!  
17 May 1996, 24 May 1996, 31 May 1996!

**SKI UTAH RENTALS**  
**WELCOMES YOU!!!**

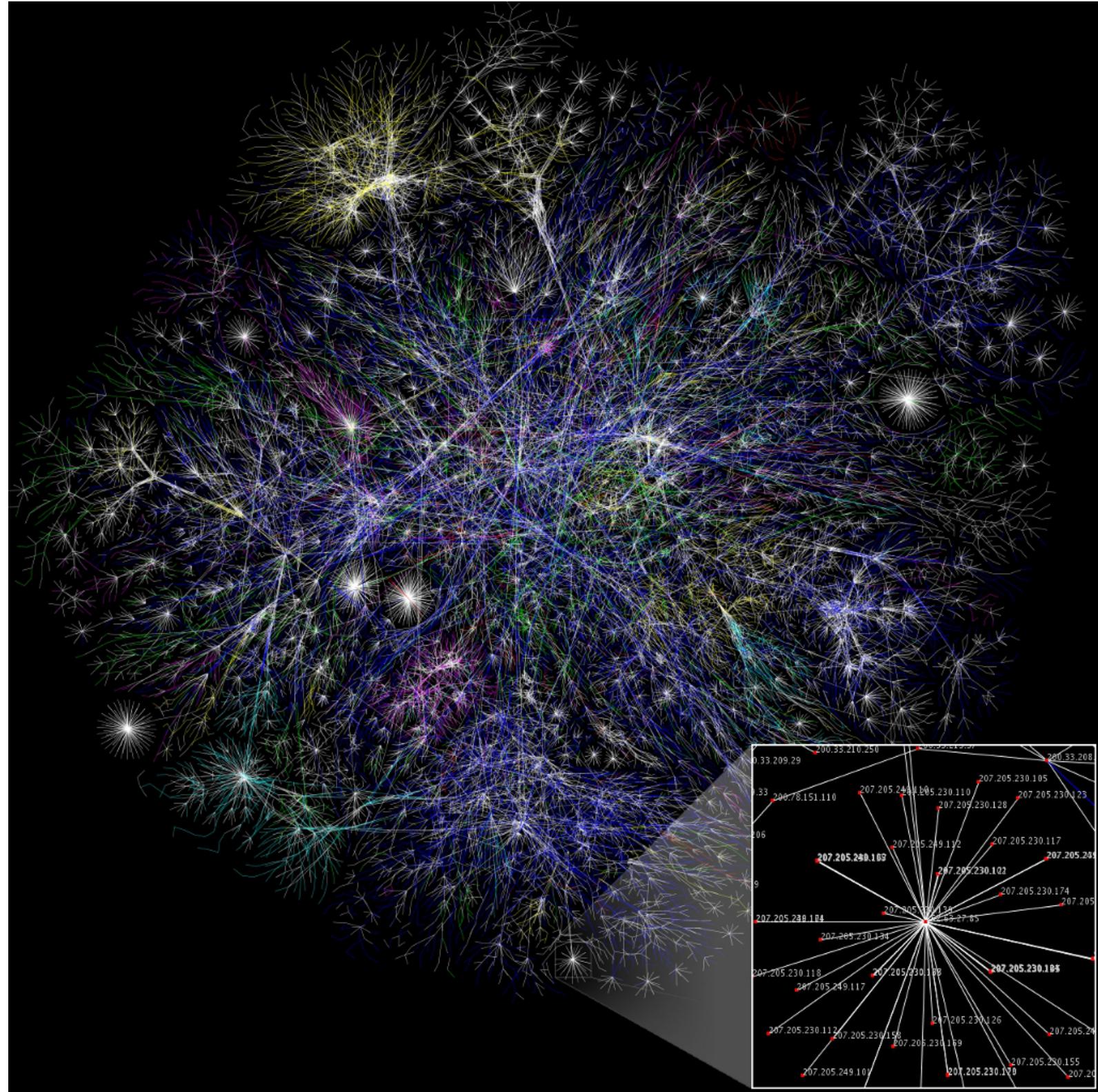
 full convention schedule

**SKI**  
**ALTA, SNOWBIRD, SOLITUDE & BRIGHTON**

**HOT RATES**



# Internet scale



# Google (1998)



search “google in 1998”

# 1997

D

Pro

Idea: A



## The PageRank Citation Ranking: Bringing Order to the Web

January 29, 1998

### Abstract

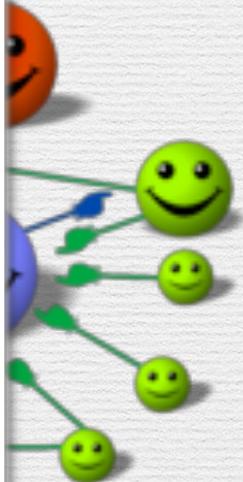
The importance of a Web page is an inherently subjective matter, which depends on the readers interests, knowledge and attitudes. But there is still much that can be said objectively about the relative importance of Web pages. This paper describes PageRank, a method for rating Web pages objectively and mechanically, effectively measuring the human interest and attention devoted to them.

We compare PageRank to an idealized random Web surfer. We show how to efficiently compute PageRank for large numbers of pages. And, we show how to apply PageRank to search and to user navigation.

### 1 Introduction and Motivation

The World Wide Web creates many new challenges for information retrieval. It is very large and heterogeneous. Current estimates are that there are over 150 million web pages with a doubling life of less than one year. More importantly, the web pages are extremely diverse, ranging from "What is Joe having for lunch today?" to journals about information retrieval. In addition to these major challenges, search engines on the Web must also contend with inexperienced users and pages engineered to manipulate search engine ranking functions.

However, unlike "flat" document collections, the World Wide Web is hypertext and provides considerable auxiliary information on top of the text of the web pages, such as link structure and



! A

Get Google! updates monthly:

your e-mail

Subscribe Archive

# Larry Page and Sergey Brin

Copyright ©1998 Google Inc.

# Google

**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](#)  
http://www.stanford.edu/  
74.79% 4K - 2591993 - 01/03/97

[Stanford University Portfolio Collection](#)  
http://www.stanford.edu/home/administration/portfolio.html  
65.78% 3K - 2591993 - 01/03/97

[University of Illinois at Urbana-Champaign](#)  
http://www.uiuc.edu/  
73.26% 13K - 13/30/95 - 01/03/97

[Indiana University](#)  
http://www.indiana.edu/  
68.38% 2K - 09/23/95 - 01/03/97

[University of California, Irvine](#)  
http://www.uci.edu/  
68.07% 3K - 13/30/95 - 01/03/97

[University of Minnesota](#)  
http://www.umn.edu/  
67.05% 0K - 13/26/95 - 01/03/97

[Iowa State University Homepage](#)  
http://www.iastate.edu/  
66.66% 3K - 13/28/95 - 01/03/97

[The University of Michigan](#)  
http://www.umich.edu/  
66.35% 2K - 2591993 - 01/03/97

[Mississippi State University](#)  
http://www.msstate.edu/  
66.35% 3K - 2591993 - 01/03/97

[Northwestern University: NUInfo](#)  
http://www.nwu.edu/  
66.15% 3K - 13/24/95 - 01/03/97

[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.oregonu.edu/> - size 1K - 16 Dec 95

[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 95

[Keio University Shonan Fujisawa Campus \(SFC\)](#)  
B\$3\$N%Z!EFnF#Bt%-%c%Q%99 (B(SFC) \$B\$N (B WWW \$B%\$BCmOU=q\$- (B \$B\$bFI\$s\$G\$/@\$5\$\$#! (B. Nihongo | English. SFC \$B>pJs (B. [ \$B%a%G%#%"%;%e%?!\*...  
<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.mnus.edu/> - size 3K - 27 Nov 95

[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](#)

next 10 ← ↑ ↓ →

# Background

- Early search engines mainly compare content similarity of the query and the indexed pages. I.e.,
  - They use information retrieval methods, cosine, TF-IDF, ...
- From 1996, it became clear that content similarity alone was no longer sufficient.
  - The number of pages grew rapidly in the mid-late 1990's.
    - Try “classification technique”, Google estimates: 10 million relevant pages.
    - How to choose only 30-40 pages and rank them suitably to present to the user?
  - Content similarity is easily spammed.
    - A page owner can repeat some words and add many related words to boost the rankings of his pages and/or to make the pages relevant to a large number of queries.

I am text. I am link. Can you see me?

I am text. I am link. Can you see me?

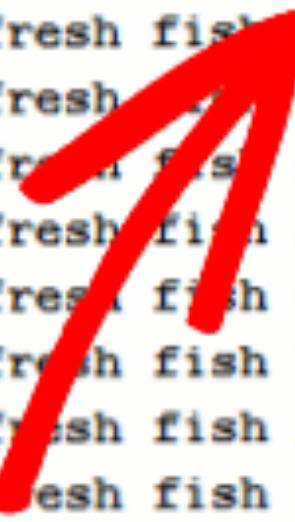
I am text. I am link. Can you see me?

I am text. I am link. Can you see me?

I am text. I am link. Can you see me?

I am text. I am link. Can you see me?

```
1 wholesale fishmonger fresh fish frozen fish surrey
2 wholesale fishmonger fresh fish frozen fish sussex
3 wholesale fishmonger fresh fish frozen fish Hastings
4 wholesale fishmonger fresh fish frozen fish Arundel
5 wholesale fishmonger fresh fish frozen fish Brighton
6 wholesale fishmonger fresh fish frozen fish Battle
7 wholesale fishmonger fresh fish frozen fish Bells Yew Green
8 wholesale fishmonger fresh fish frozen fish Billingshurst
9 wholesale fishmonger fresh fish frozen fish Blackboys
```



# Background

- Starting around 1996, researchers began to work on the problem. They resort to **hyperlinks**.
  - In Feb, 1997, Yanhong Li (Robin Li), Scotch Plains, NJ, filed a hyperlink based search patent. The method uses words in anchor text of hyperlinks.
- Web pages on the other hand are connected through hyperlinks, which carry important information.
  - **Some hyperlinks**: organize information at the same site.
  - **Other hyperlinks**: point to pages from other Web sites. Such out-going hyperlinks often indicate an **implicit conveyance of authority** to the pages being pointed to.
- Those pages that are pointed to by many other pages are likely to contain authoritative information.

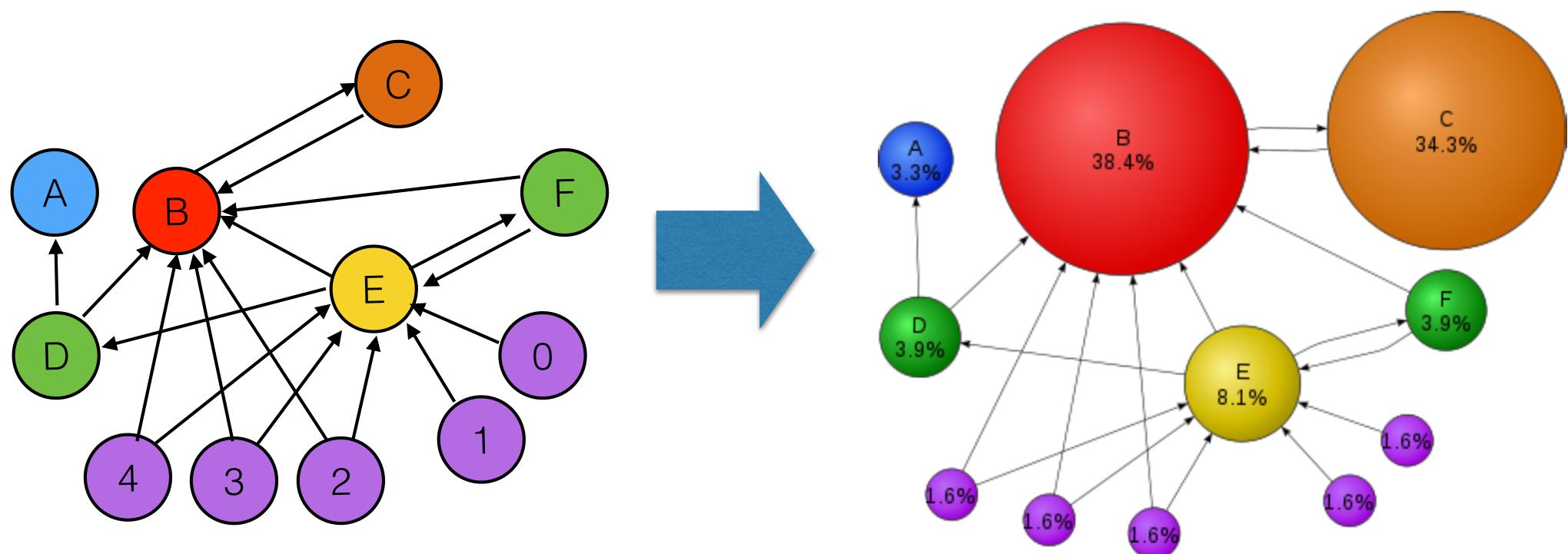
- Starting around 1996, researchers began to work on the problem. They resort to **hyperlinks**.
  - In Feb, 1997, Yanhong Li (Robin Li), Scotch Plains, NJ, filed a hyperlink based search patent. The method uses words in anchor text of hyperlinks.



# Background

- During 1997-1998, two most influential hyperlink based search algorithms **PageRank** and **HITS** were reported.
- Both algorithms are related to **social networks**. They exploit the hyperlinks of the Web to rank pages according to their levels of “prestige” or “authority”.
  - **HITS**: Jon Kleinberg (Cornel University), at *Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, January 1998
  - **PageRank**: Sergey Brin and Larry Page, PhD students from Stanford University, at *Seventh International World Wide Web Conference (WWW7)* in April, 1998.
- **PageRank powers the Google search engine.**

# Graph Ranking: How?



# 1998

D

Pro

Idea: A



## The PageRank Citation Ranking: Bringing Order to the Web

January 29, 1998

### Abstract

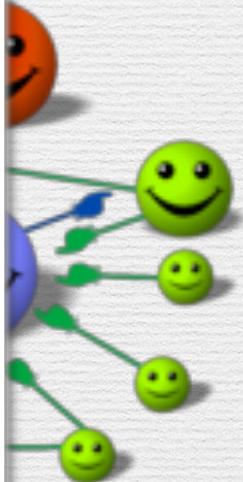
The importance of a Web page is an inherently subjective matter, which depends on the readers interests, knowledge and attitudes. But there is still much that can be said objectively about the relative importance of Web pages. This paper describes PageRank, a method for rating Web pages objectively and mechanically, effectively measuring the human interest and attention devoted to them.

We compare PageRank to an idealized random Web surfer. We show how to efficiently compute PageRank for large numbers of pages. And, we show how to apply PageRank to search and to user navigation.

### 1 Introduction and Motivation

The World Wide Web creates many new challenges for information retrieval. It is very large and heterogeneous. Current estimates are that there are over 150 million web pages with a doubling life of less than one year. More importantly, the web pages are extremely diverse, ranging from "What is Joe having for lunch today?" to journals about information retrieval. In addition to these major challenges, search engines on the Web must also contend with inexperienced users and pages engineered to manipulate search engine ranking functions.

However, unlike "flat" document collections, the World Wide Web is hypertext and provides considerable auxiliary information on top of the text of the web pages, such as link structure and



! A

Get Google! updates monthly:

your e-mail

Subscribe Archive

# Larry Page and Sergey Brin

Copyright ©1998 Google Inc.

1997

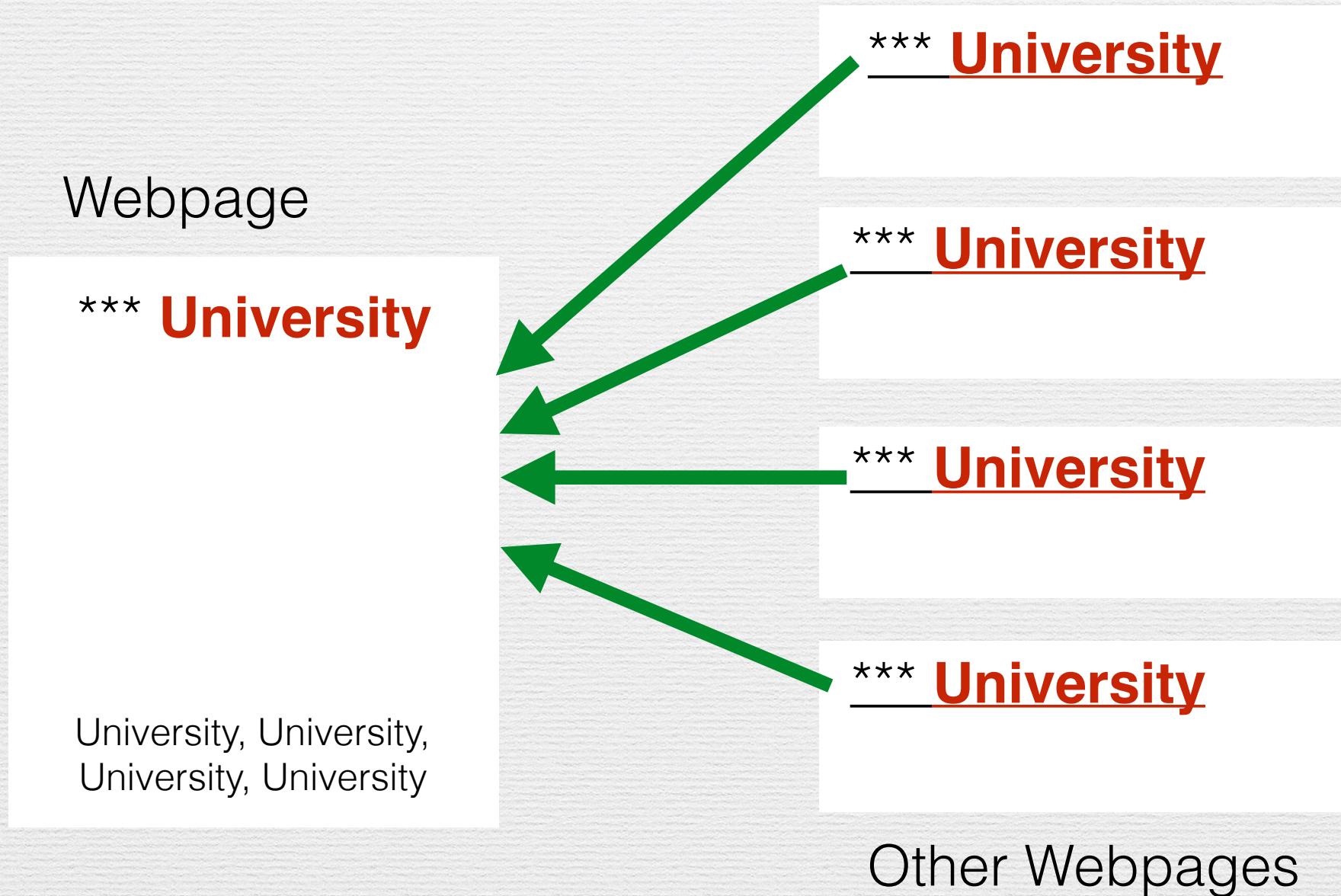


Counting Anchor-Text

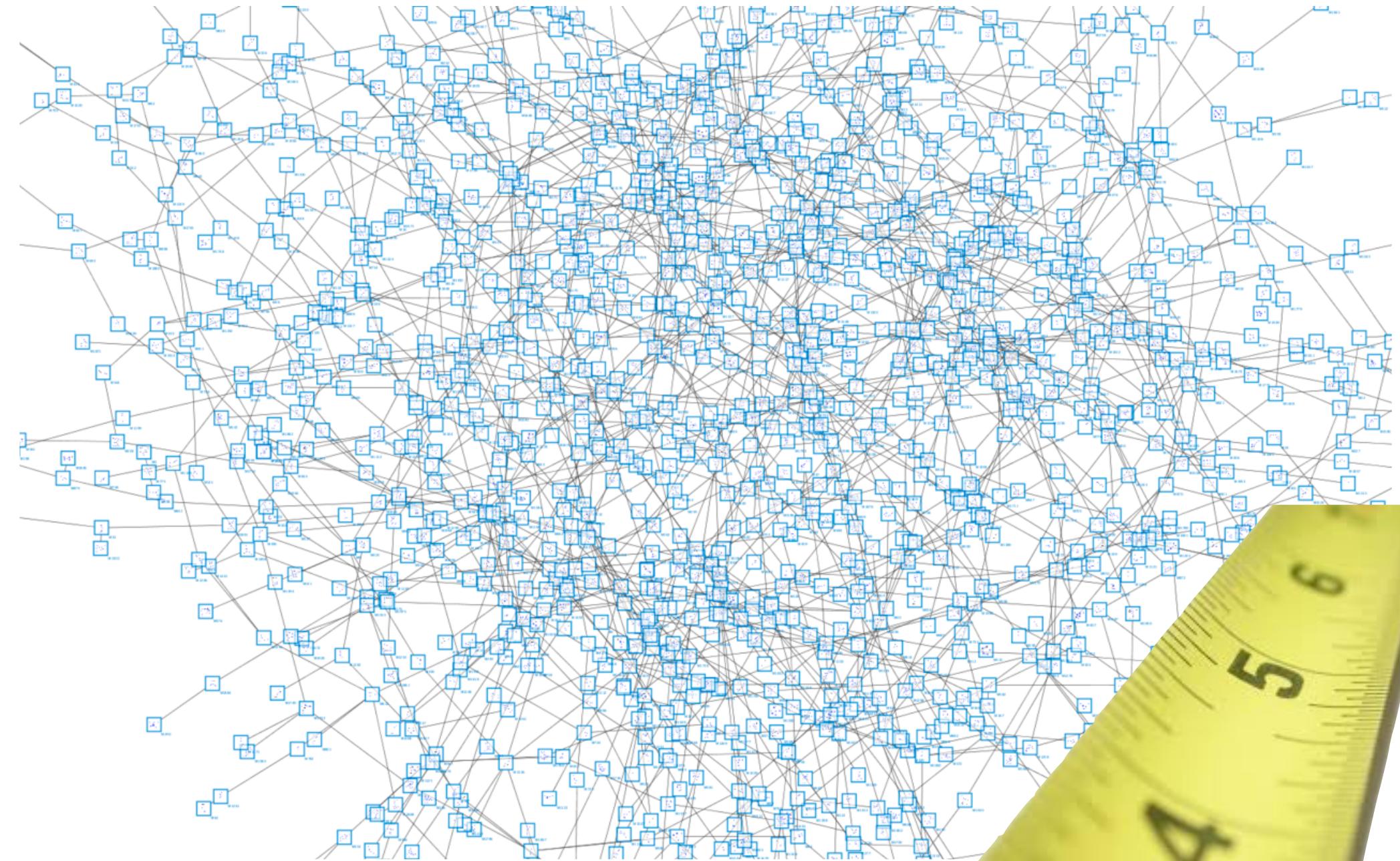


Counting In-Links

USPTO, "Hypertext Document Retrieval System and Method", US Patent number: 5920859, Inventor: **Yanhong Li**, Filing date: **5 February 1997**, Issue date: 6 July 1999

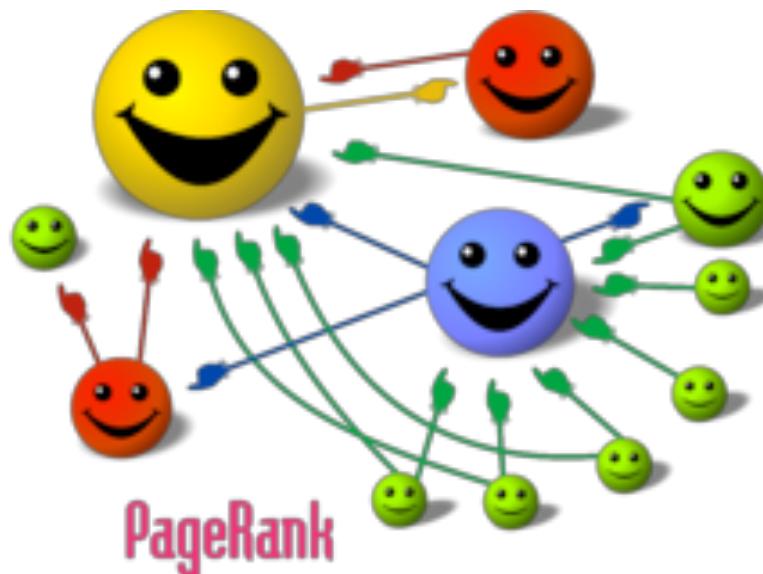


# Evaluating Graph Nodes



# PageRank

- Web**Page**: ranking pages
- Larry **Page**: author of the paper



# The PageRank Citation Ranking: Bringing Order to the Web

January 29, 1998

## Abstract

The importance of a Web page is an inherently subjective matter, which depends on the readers interests, knowledge and attitudes. But there is still much that can be said objectively about the relative importance of Web pages. This paper describes PageRank, a method for rating Web pages objectively and mechanically, effectively measuring the human interest and attention devoted to them.

We compare PageRank to an idealized random Web surfer. We show how to efficiently compute PageRank for large numbers of pages. And, we show how to apply PageRank to search and to user navigation.

## 1 Introduction and Motivation

The World Wide Web creates many new challenges for information retrieval. It is very large and heterogeneous. Current estimates are that there are over 150 million web pages with a doubling life of less than one year. More importantly, the web pages are extremely diverse, ranging from "What is Joe having for lunch today?" to journals about information retrieval. In addition to these major challenges, search engines on the Web must also contend with inexperienced users and pages engineered to manipulate search engine ranking functions.

However, unlike "flat" document collections, the World Wide Web is hypertext and provides considerable auxiliary information on top of the text of the web pages, such as link structure and

# Idea of PageRank

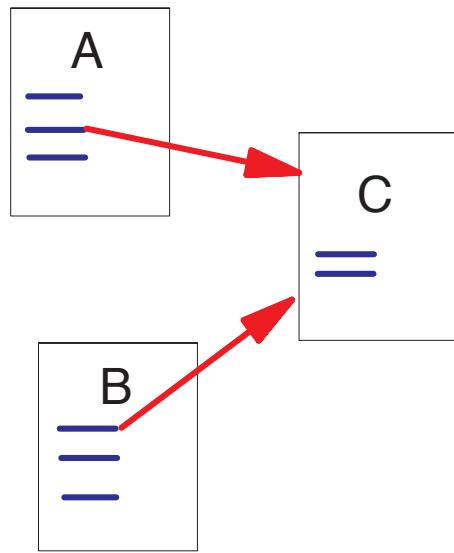


Figure 1: A and B are Backlinks of C

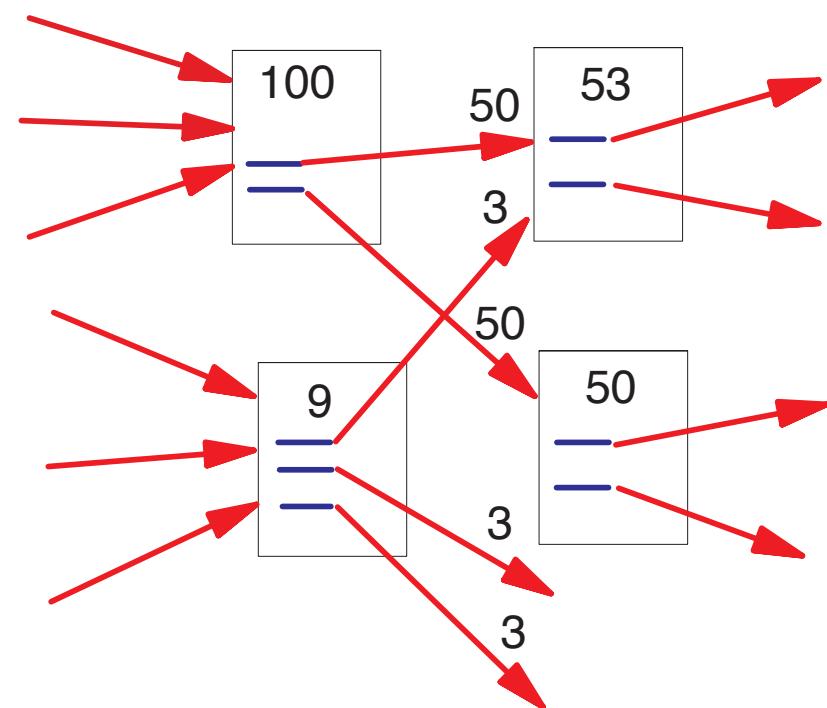


Figure 2: Simplified PageRank Calculation

# Why these ideas work

- Santa Approach (Yahoo vs Google) - Crowdsourcing

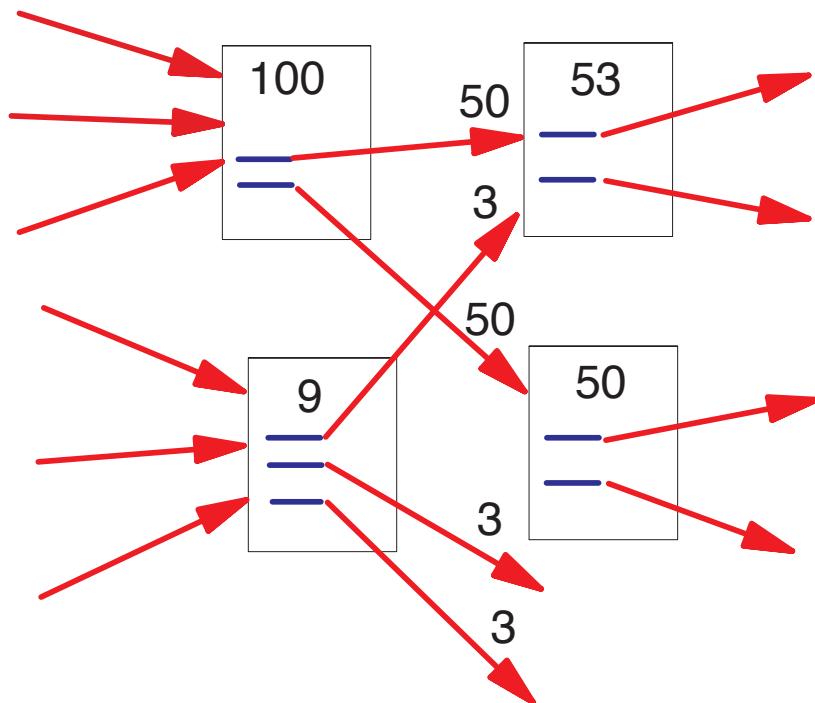
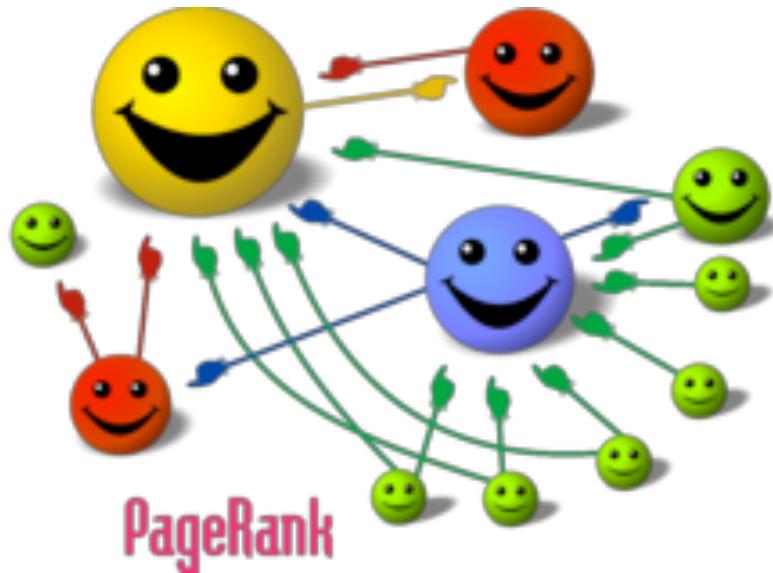


Figure 2: Simplified PageRank Calculation



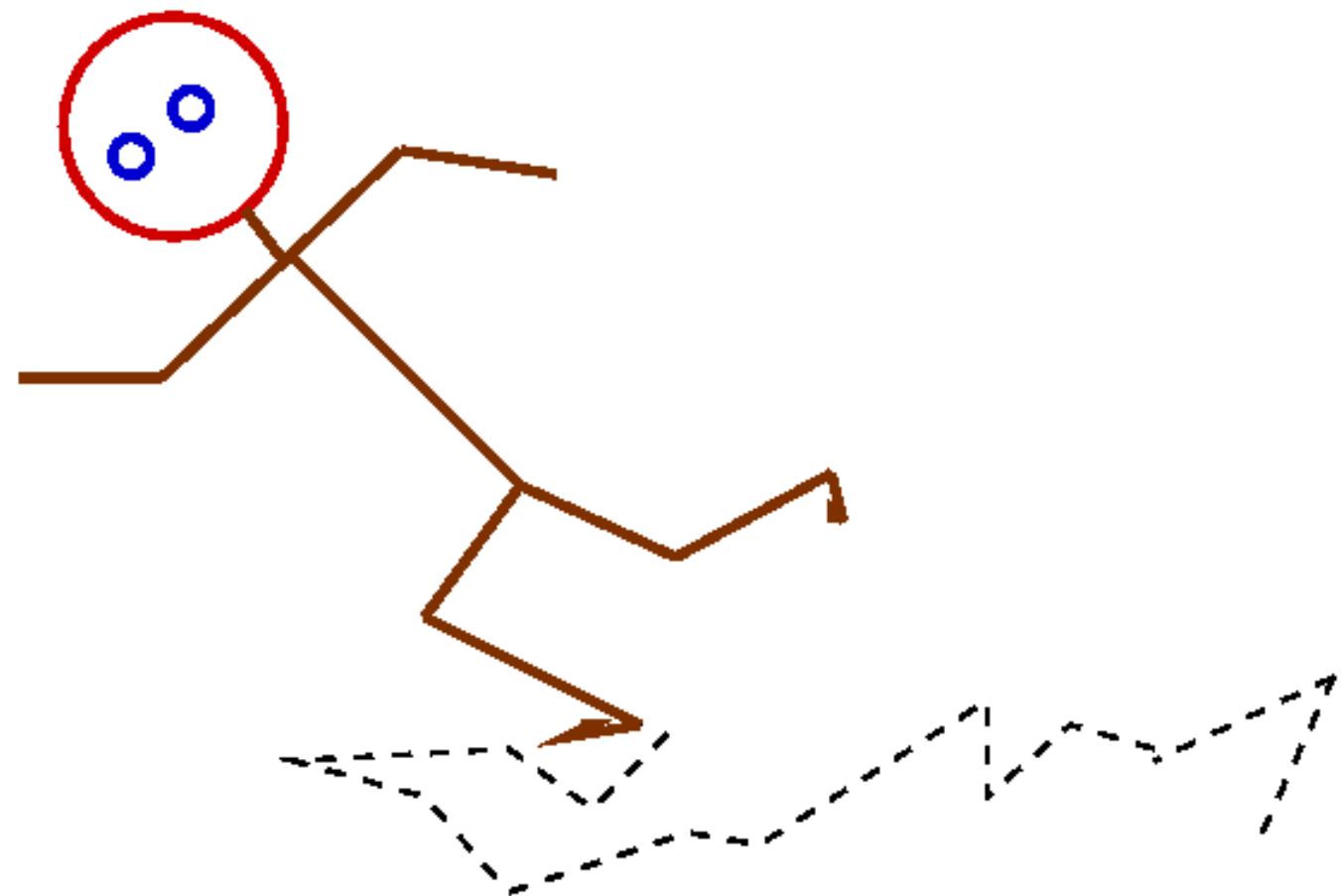
<https://youtu.be/dCOf6Su0Yvw?t=2m55s>

# PageRank

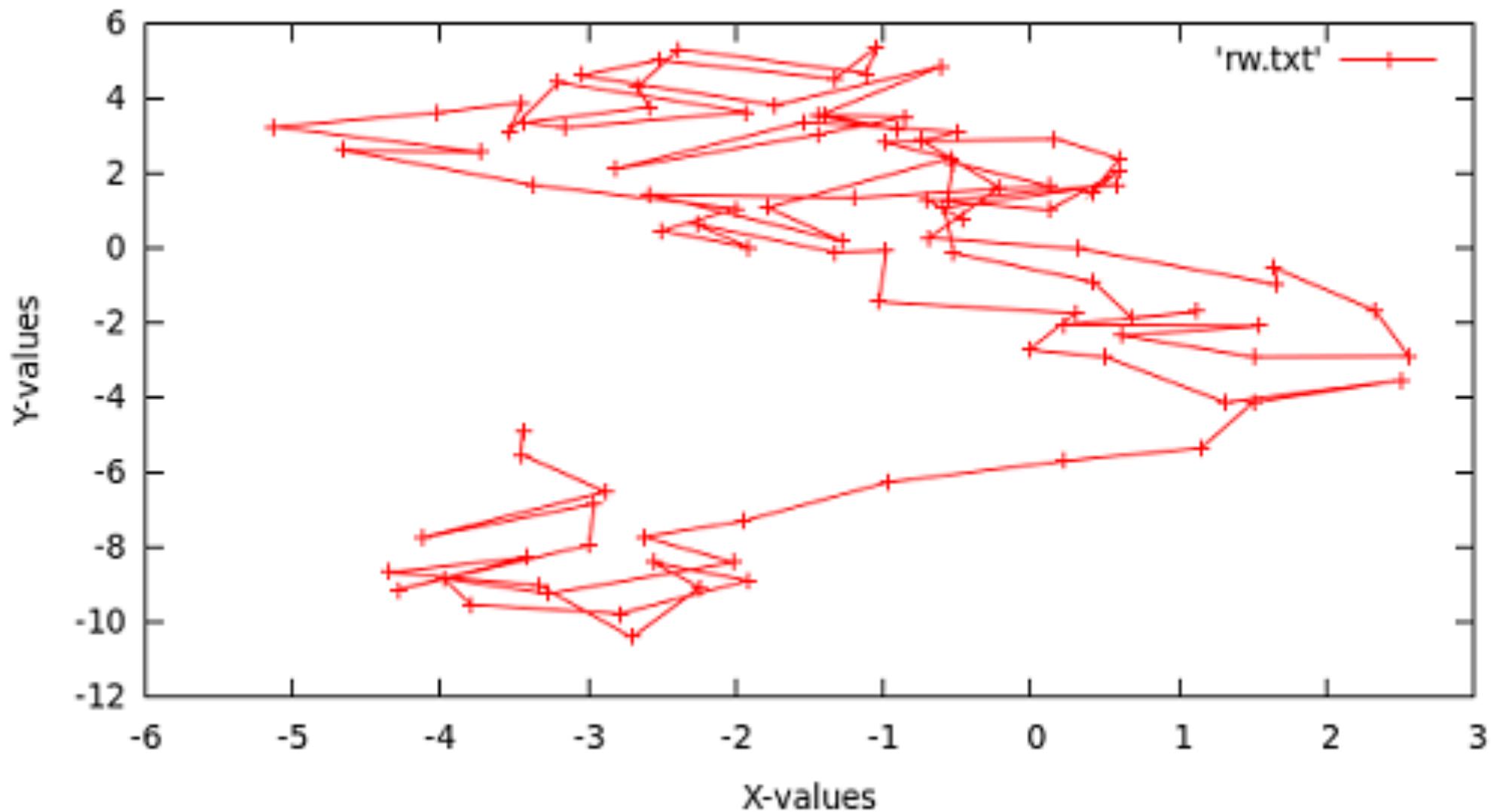


- **Important** pages should be linked by **other important** pages
  - **Chicken and Egg**

# Random Walk

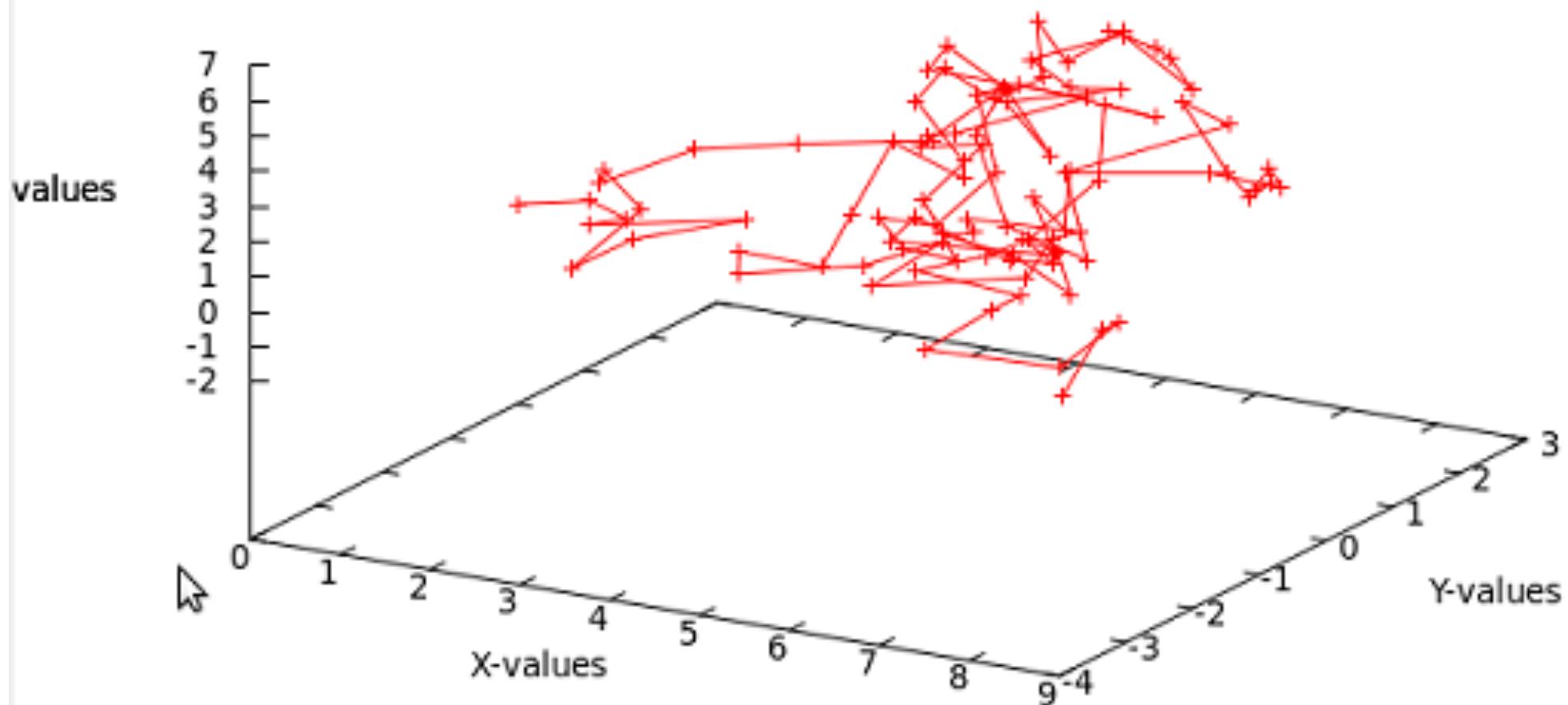


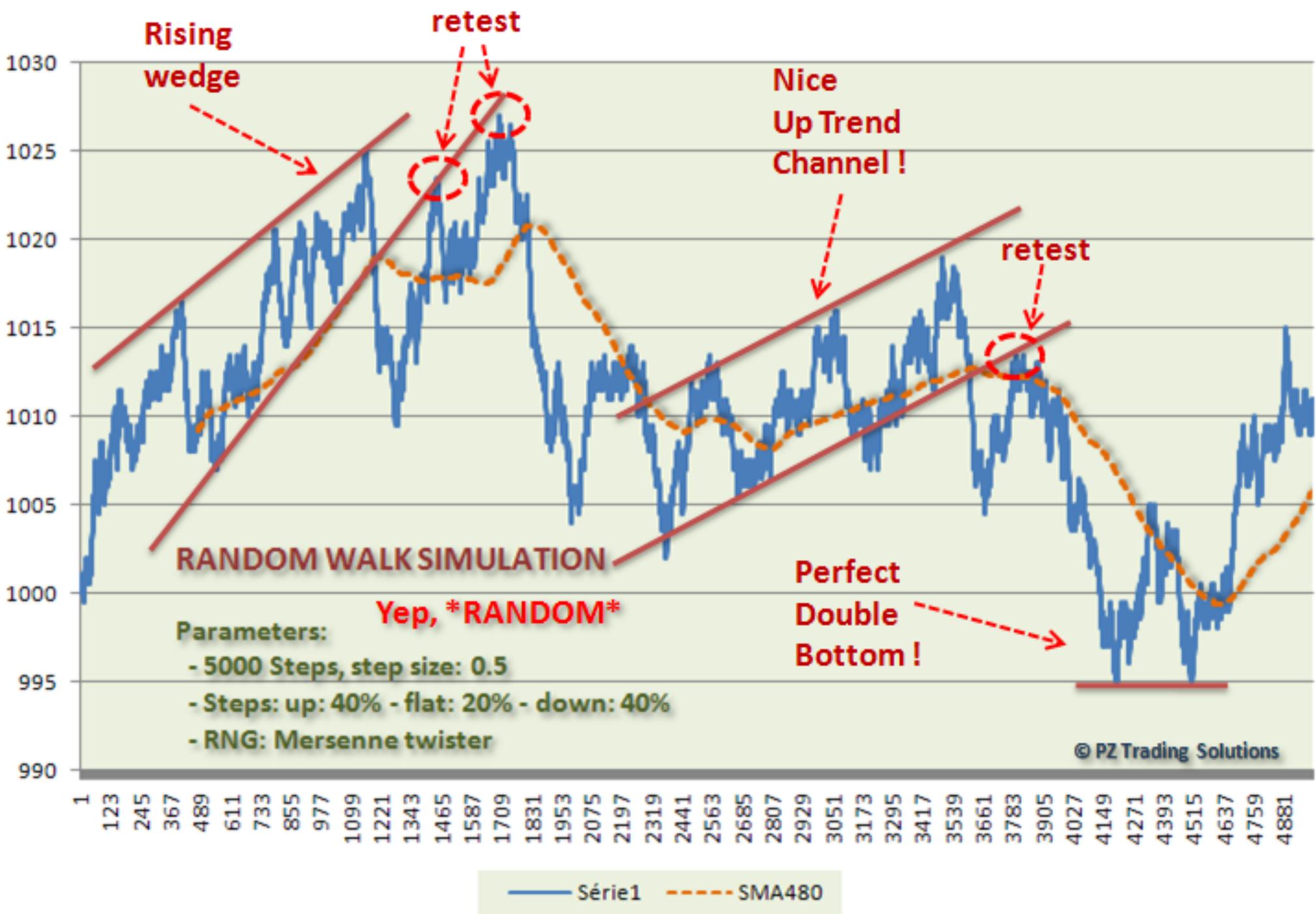
## 2D Random Walk



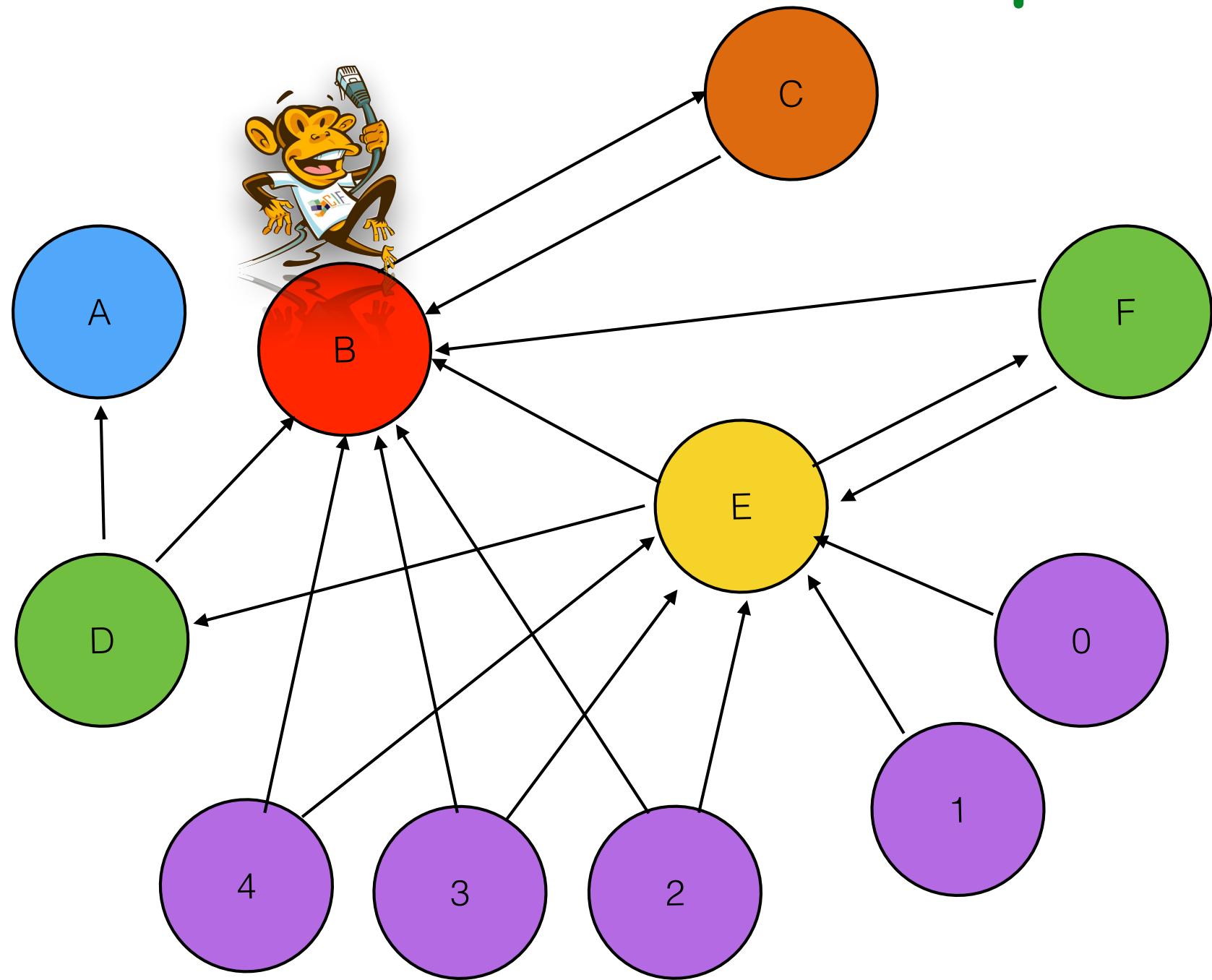
3D Random Walk (100 steps)

'rw3d.txt' —+

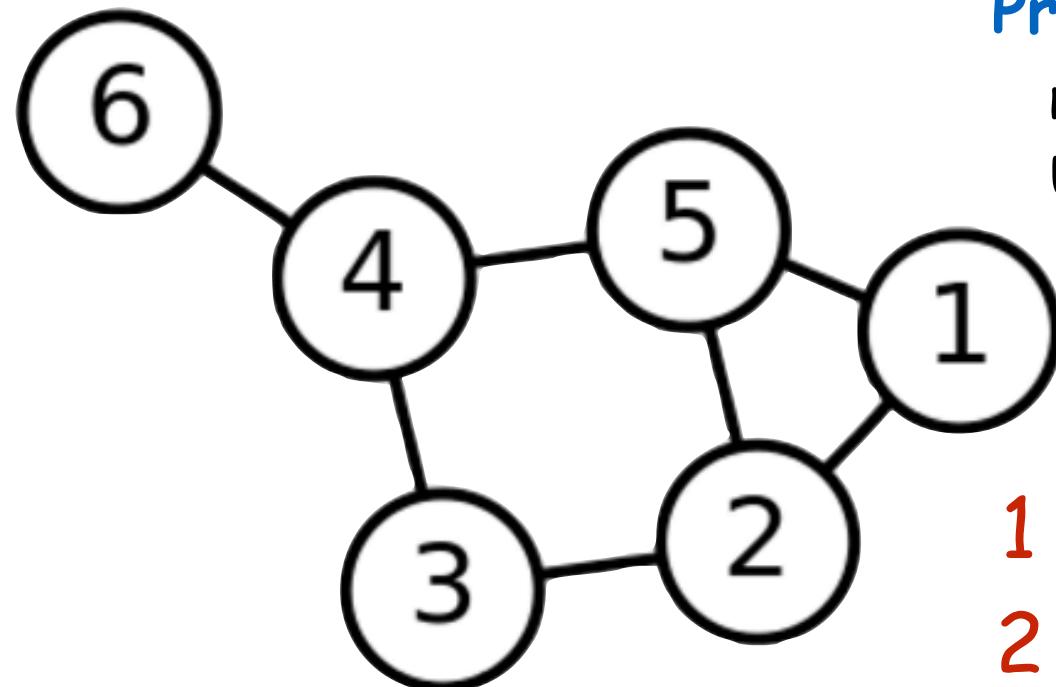




# Random Walk on Graph



# Adjacency Matrix (A)



Properties:

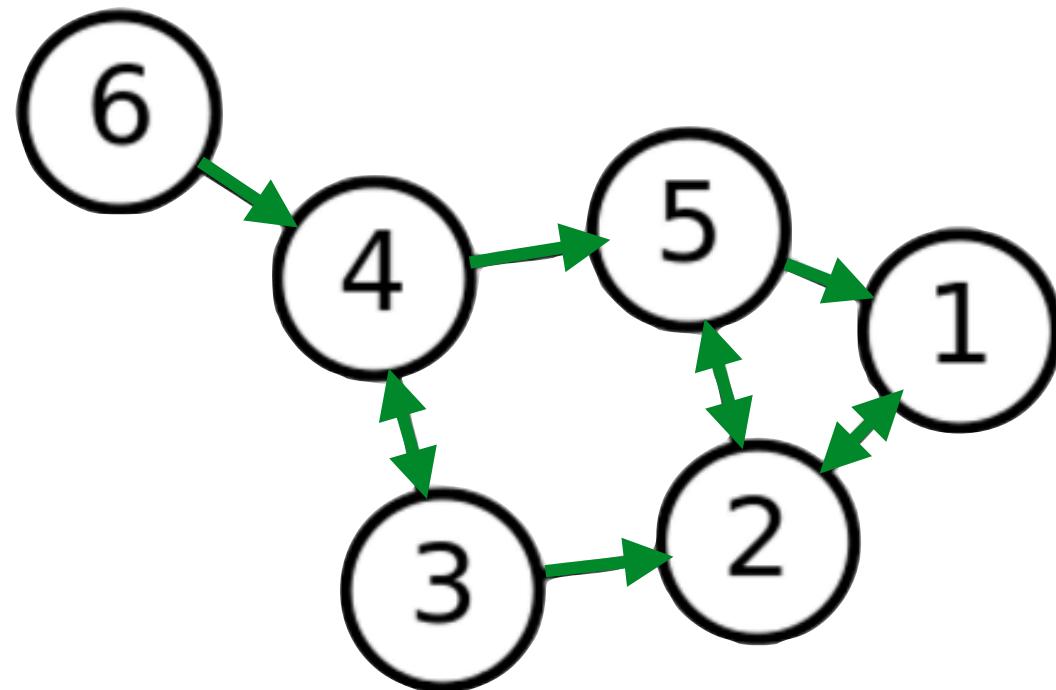
$n \times n$  matrix

Undirected graph  $\rightarrow$  Symmetric matrix

$A =$

|   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|
|   | 1 | 2 | 3 | 4 | 5 | 6 |
| 1 | 0 | 1 | 0 | 0 | 1 | 0 |
| 2 | 1 | 0 | 1 | 0 | 1 | 0 |
| 3 | 0 | 1 | 0 | 1 | 0 | 0 |
| 4 | 0 | 0 | 1 | 0 | 1 | 1 |
| 5 | 1 | 1 | 0 | 1 | 0 | 0 |
| 6 | 0 | 0 | 0 | 1 | 0 | 0 |

# Directed Graph

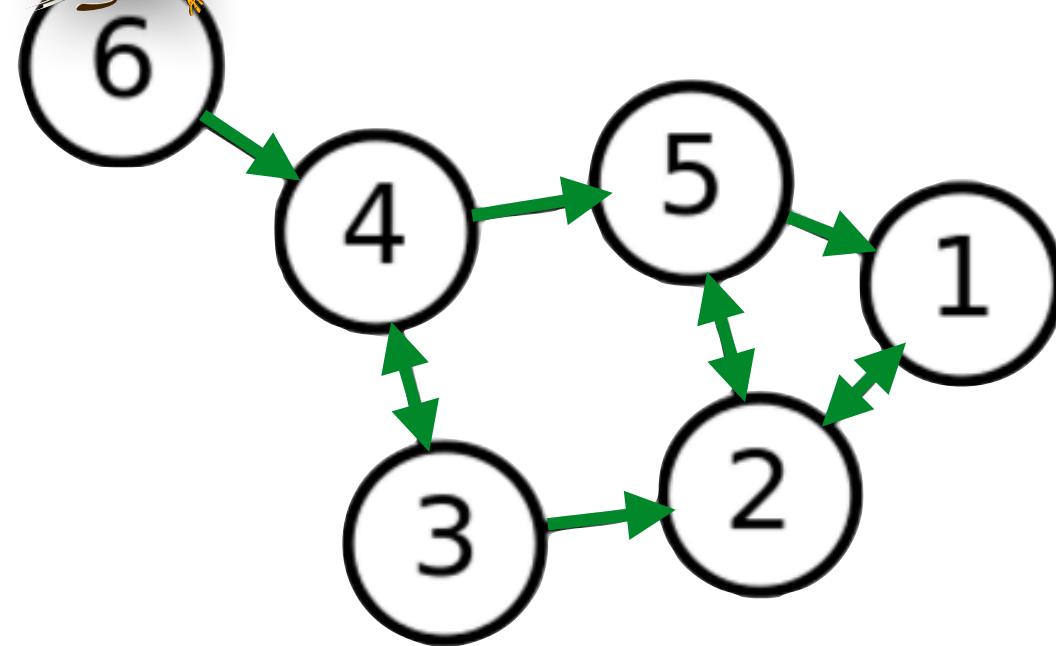


NOT Symmetric

$A =$

|   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|
|   | 1 | 2 | 3 | 4 | 5 | 6 |
| 1 | 0 | 1 | 0 | 0 | 1 | 0 |
| 2 | 1 | 0 | 1 | 0 | 1 | 0 |
| 3 | 0 | 0 | 0 | 1 | 0 | 0 |
| 4 | 0 | 0 | 1 | 0 | 0 | 1 |
| 5 | 0 | 1 | 0 | 1 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 | 0 | 0 |

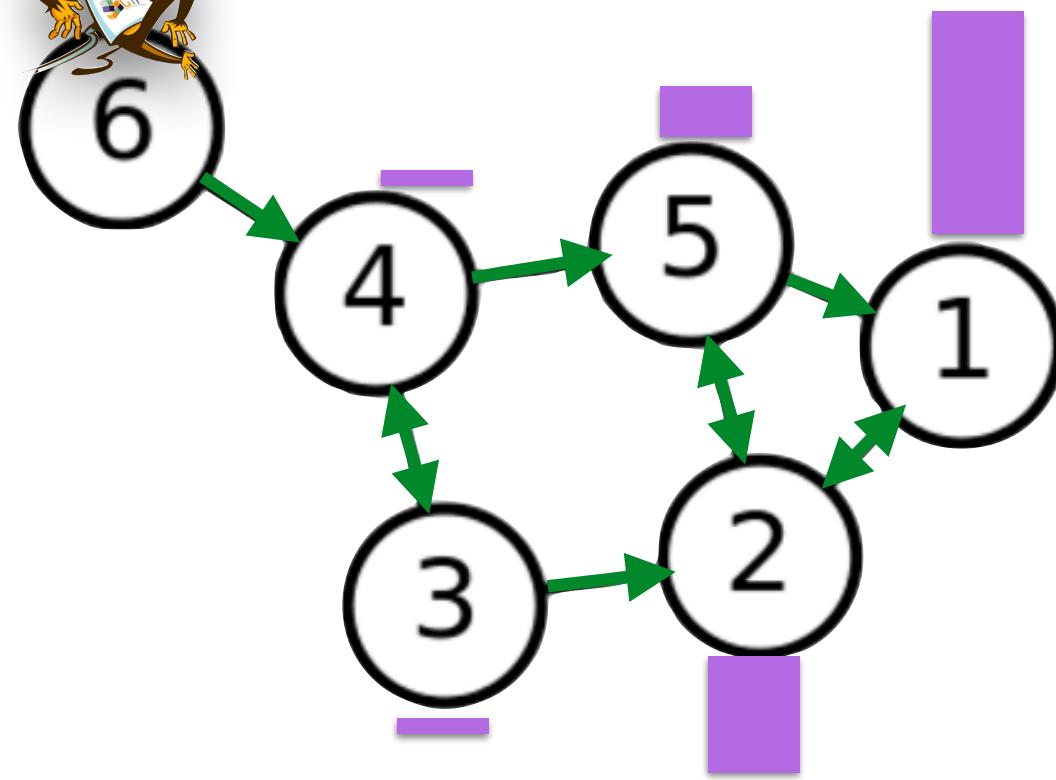
# Random Walk on Graph



$A =$

|   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|
|   | 1 | 2 | 3 | 4 | 5 | 6 |
| 1 | 0 | 1 | 0 | 0 | 1 | 0 |
| 2 | 1 | 0 | 1 | 0 | 1 | 0 |
| 3 | 0 | 0 | 0 | 1 | 0 | 0 |
| 4 | 0 | 0 | 1 | 0 | 0 | 1 |
| 5 | 0 | 1 | 0 | 1 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 | 0 | 0 |

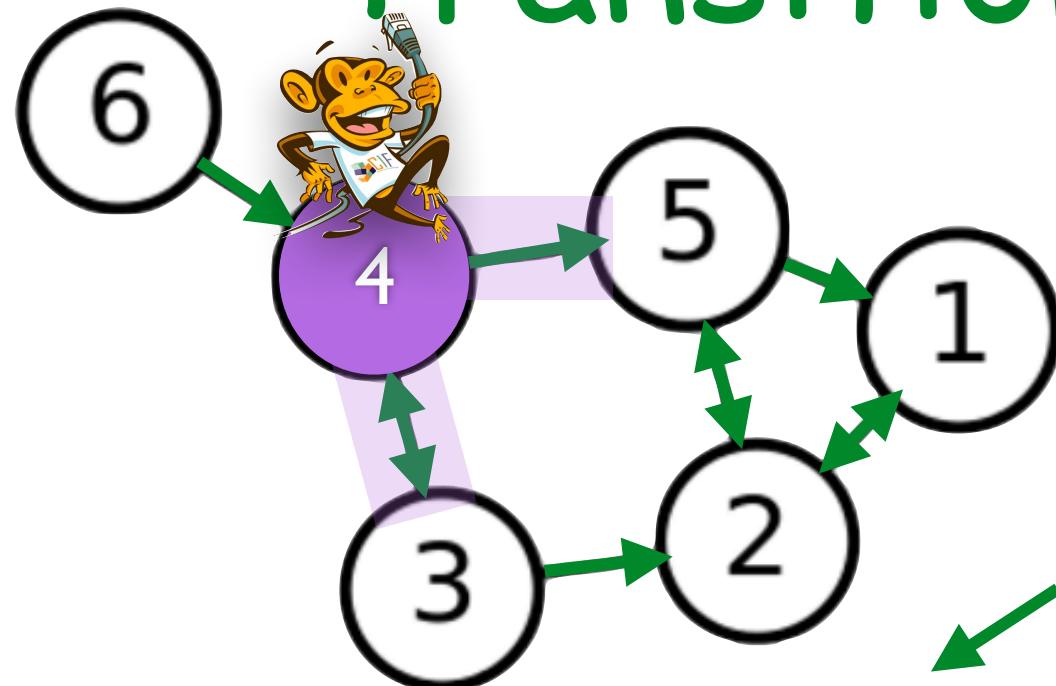
# Random Walk on Graph



$A =$

|   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 |   |
| 1 | 0 | 1 | 0 | 0 | 1 | 0 |
| 2 | 1 | 0 | 1 | 0 | 1 | 0 |
| 3 | 0 | 0 | 0 | 1 | 0 | 0 |
| 4 | 0 | 0 | 1 | 0 | 0 | 1 |
| 5 | 0 | 1 | 0 | 1 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 | 0 | 0 |

# Transition Matrix (P)



Properties:

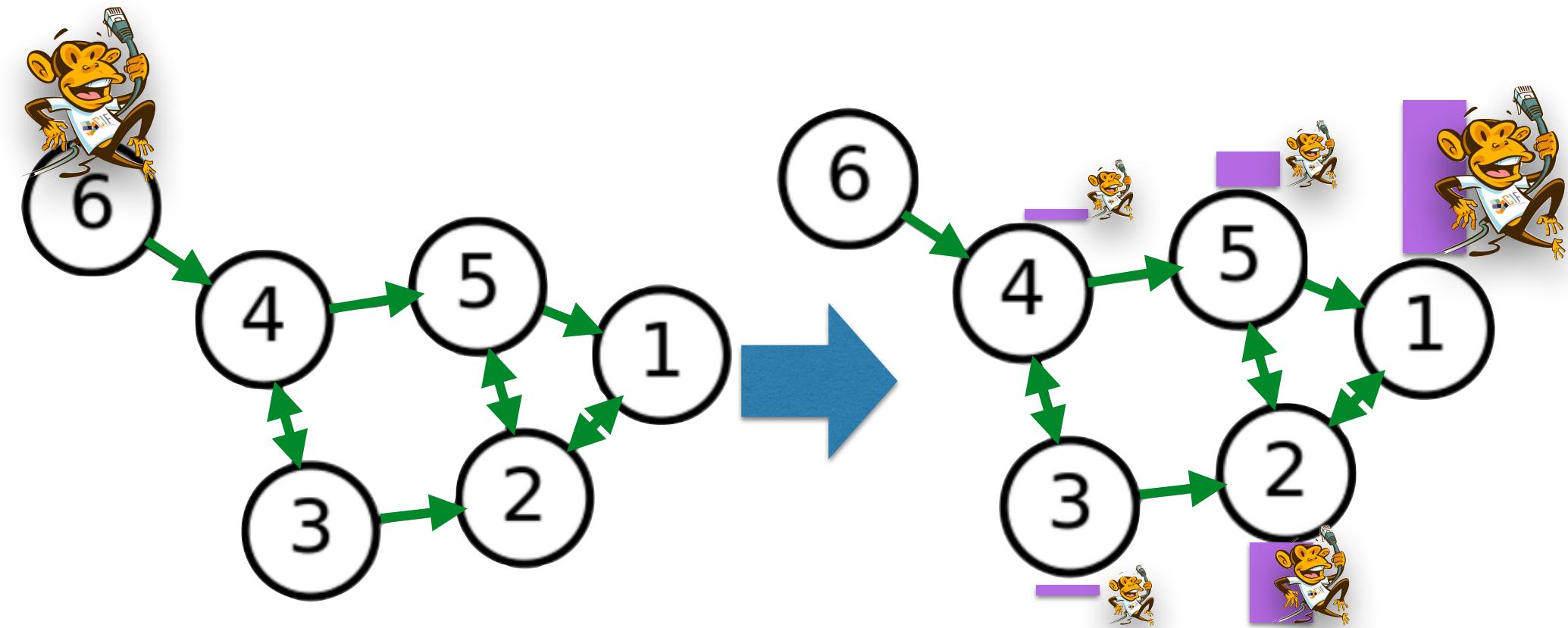
$n \times n$  matrix

normalized columns

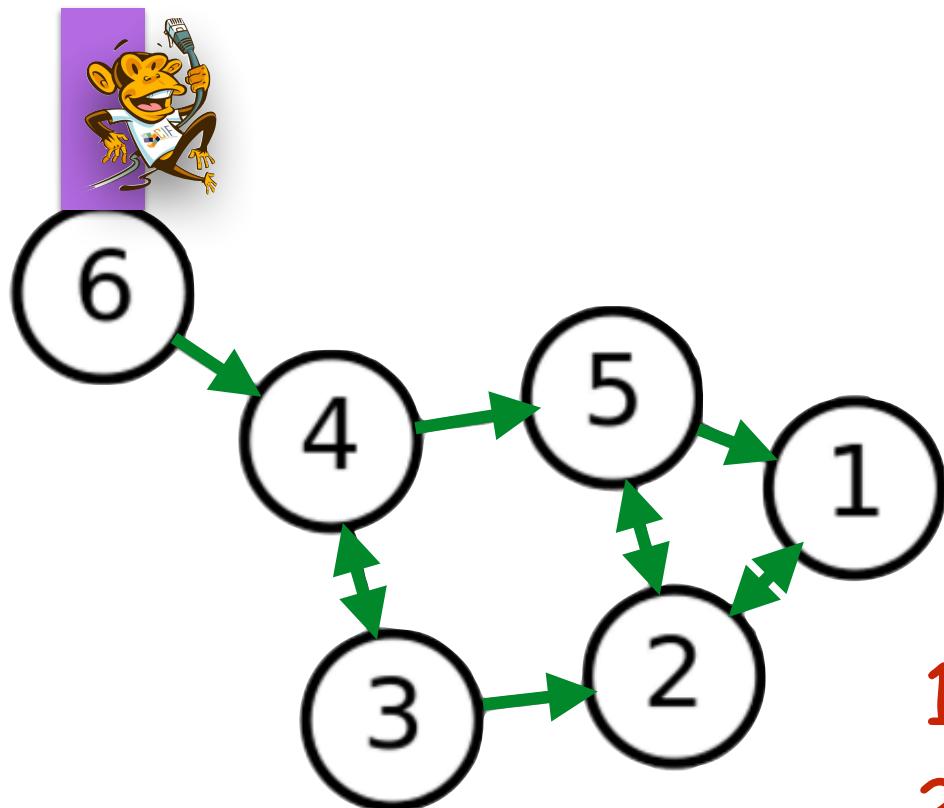
$$P =$$

|   | 1 | 2  | 3  | 4  | 5  | 6 |
|---|---|----|----|----|----|---|
| 1 | 0 | .5 | 0  | 0  | .5 | 0 |
| 2 | 1 | 0  | .5 | 0  | .5 | 0 |
| 3 | 0 | 0  | 0  | .5 | 0  | 0 |
| 4 | 0 | 0  | .5 | 0  | 0  | 1 |
| 5 | 0 | .5 | 0  | .5 | 0  | 0 |
| 6 | 0 | 0  | 0  | 0  | 0  | 0 |

# Random Walk on Graph



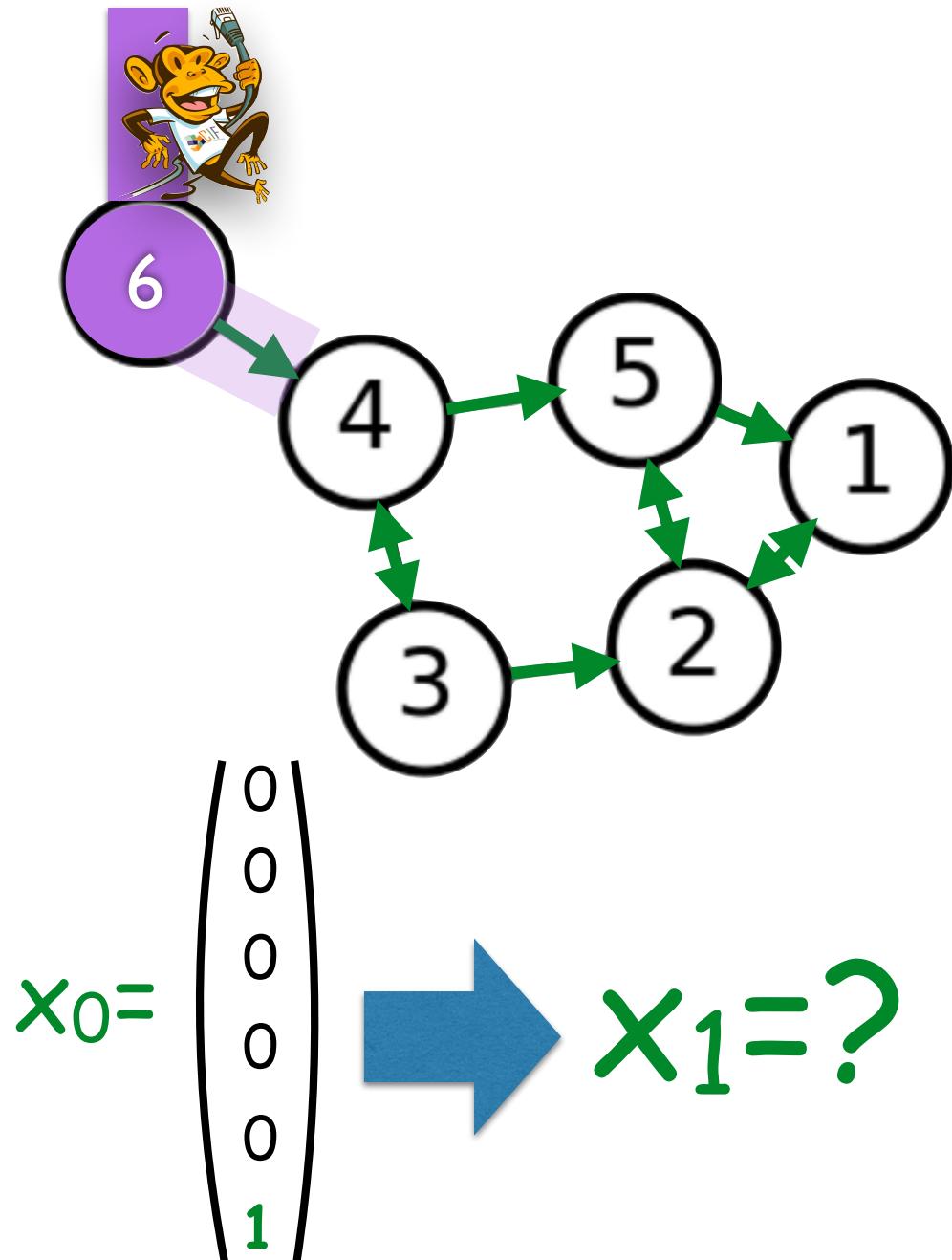
# Probability Distribution



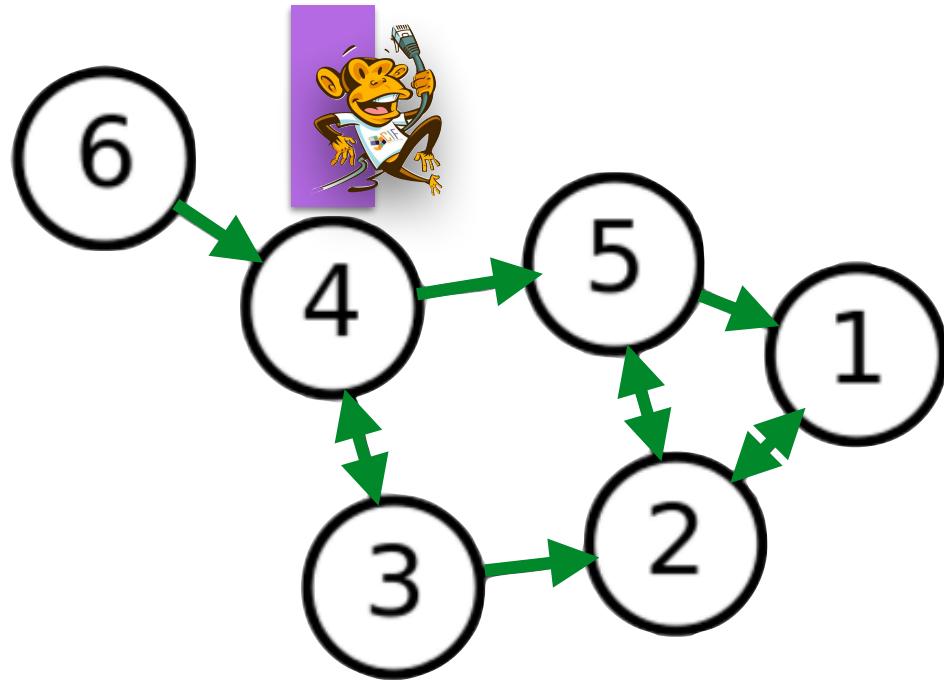
$x =$

|   |   |
|---|---|
| 1 | 0 |
| 2 | 0 |
| 3 | 0 |
| 4 | 0 |
| 5 | 0 |
| 6 | 1 |

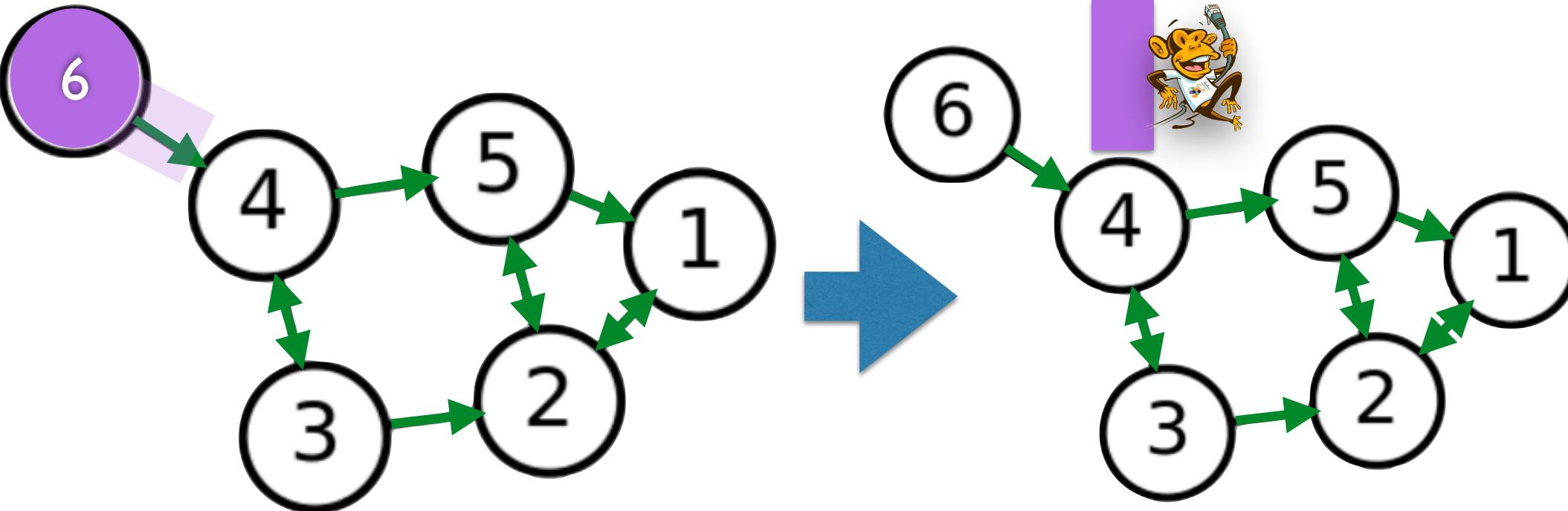
# 1 step Random Walk



# 1 step Random Walk

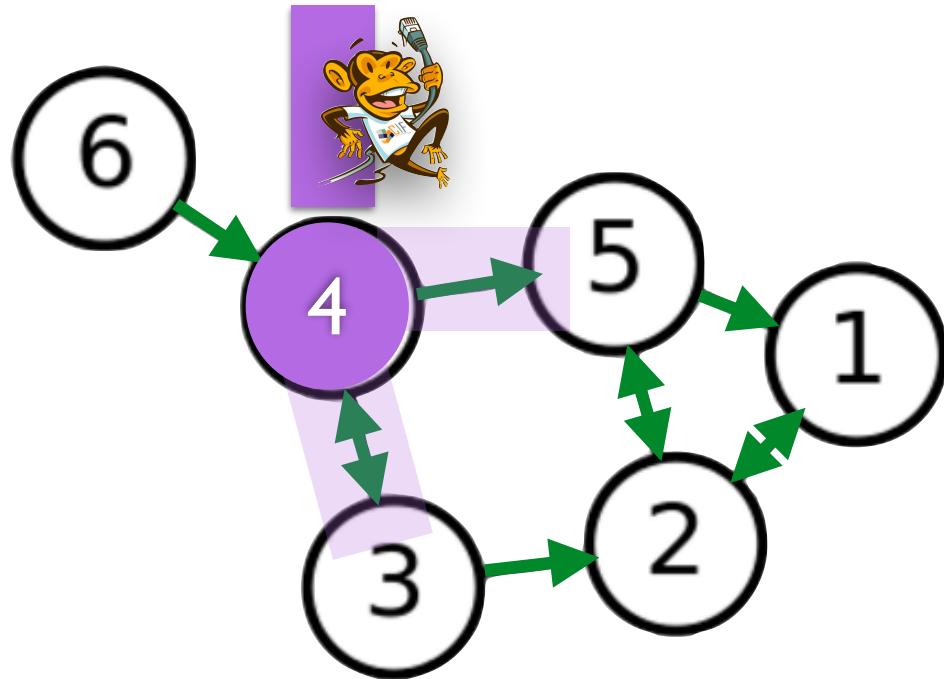


$$x_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad \xrightarrow{\text{Blue Arrow}} \quad x_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

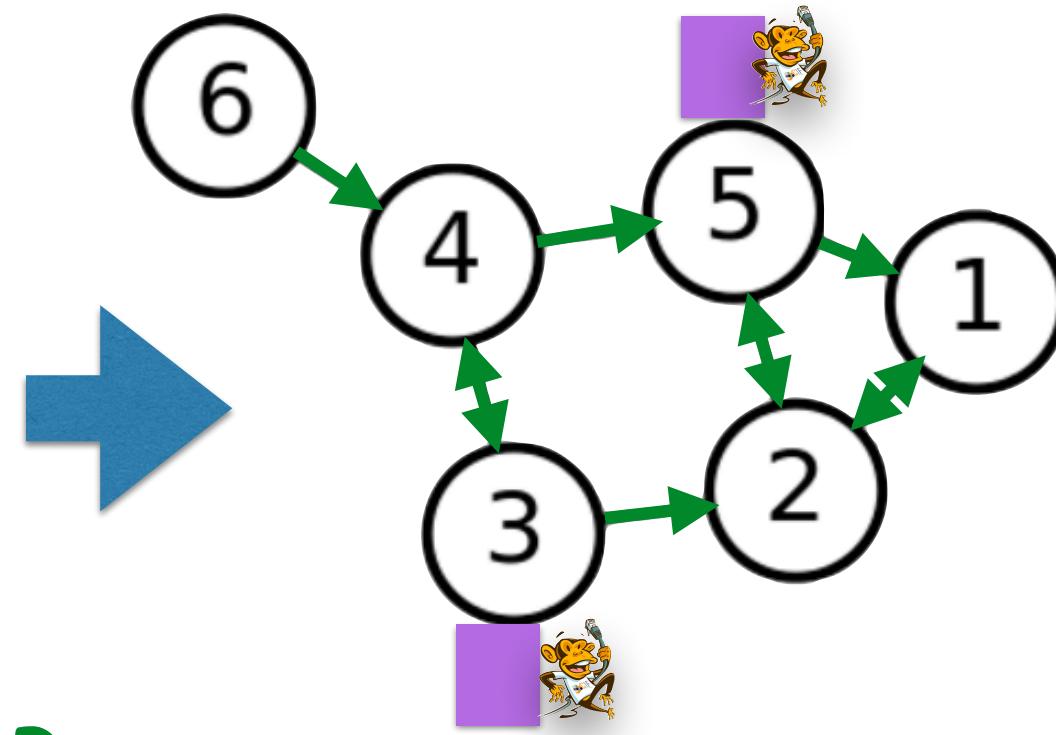
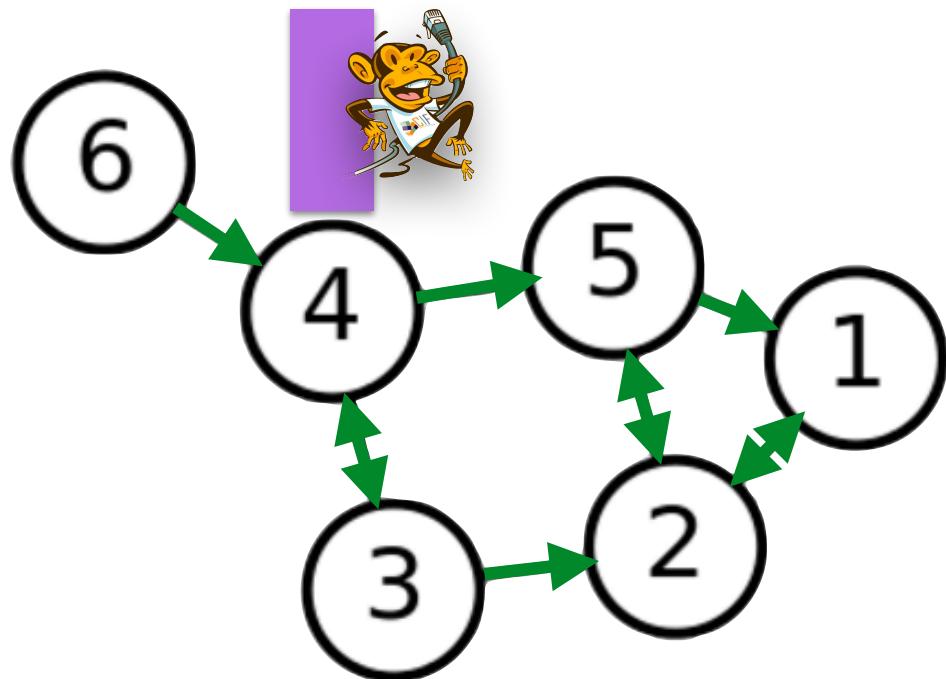


$$\begin{matrix}
 & & P & & \\
 & x_1 & & & x_0 \\
 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} & = & \begin{pmatrix} 0 & .5 & 0 & 0 & .5 & 0 \\ 1 & 0 & .5 & 0 & .5 & 0 \\ 0 & 0 & 0 & .5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & .5 & 0 & 0 & .5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} & \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} & x
 \end{pmatrix}$$

# 1 step Random Walk



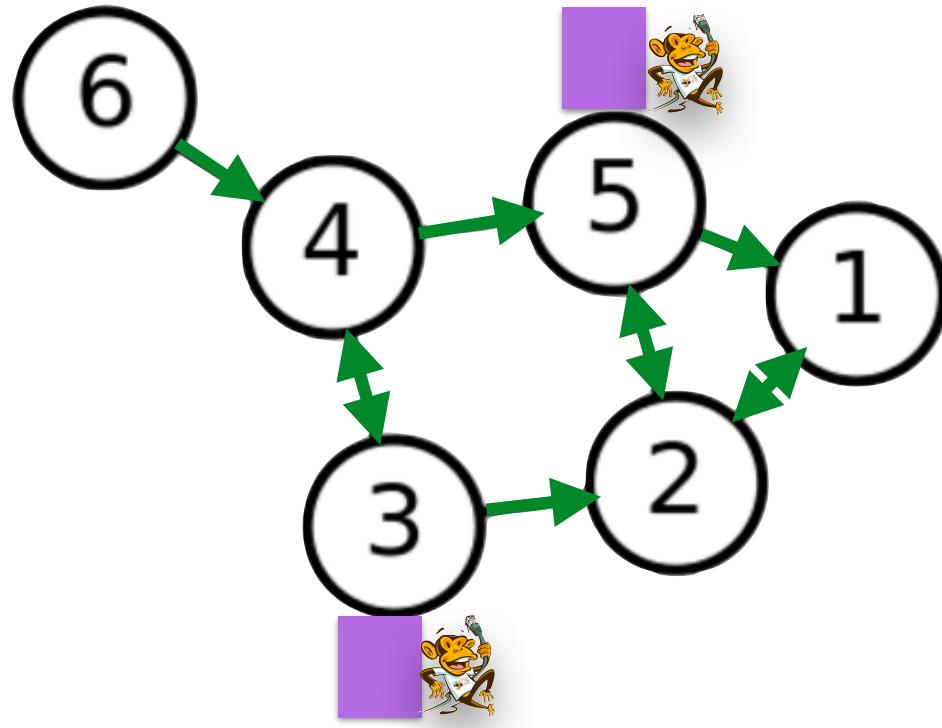
$$x_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \rightarrow x_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \rightarrow x_2 = \begin{pmatrix} 0 \\ 0 \\ .5 \\ 0 \\ .5 \\ 0 \end{pmatrix}$$



$$\begin{matrix}
 & x_2 \\
 \left[ \begin{array}{c} 0 \\ 0 \\ .5 \\ 0 \\ .5 \\ 0 \end{array} \right] & = & P & \left[ \begin{array}{c} x_1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{array} \right]
 \end{matrix}$$

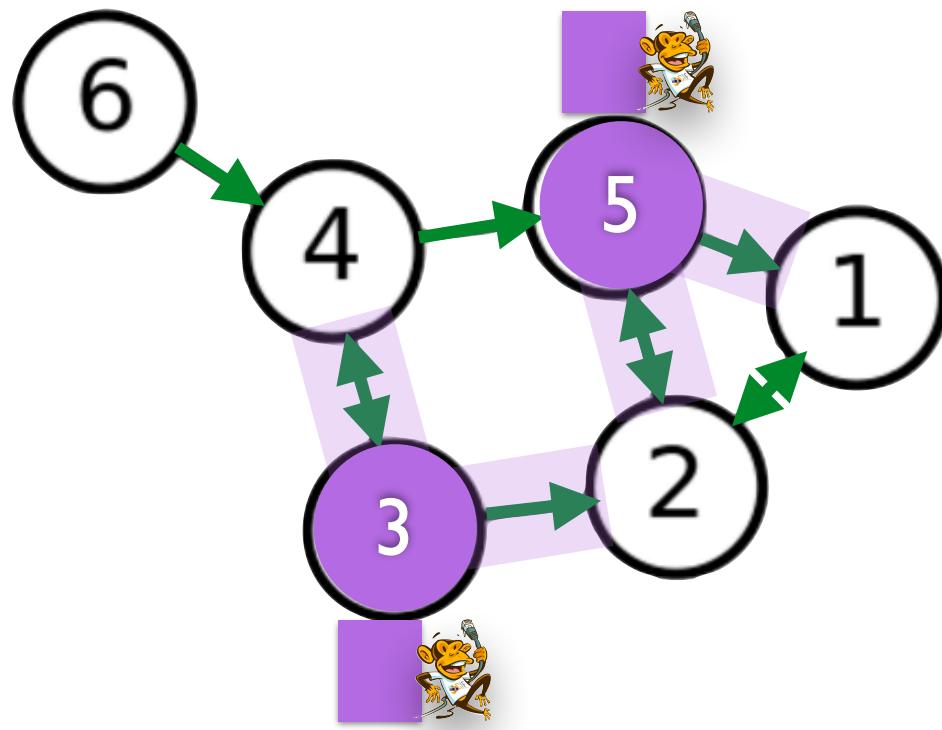
$$P = \begin{pmatrix}
 0 & .5 & 0 & 0 & .5 & 0 \\
 1 & 0 & .5 & 0 & .5 & 0 \\
 0 & 0 & 0 & .5 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 1 \\
 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0
 \end{pmatrix}$$

# 1 step Random Walk



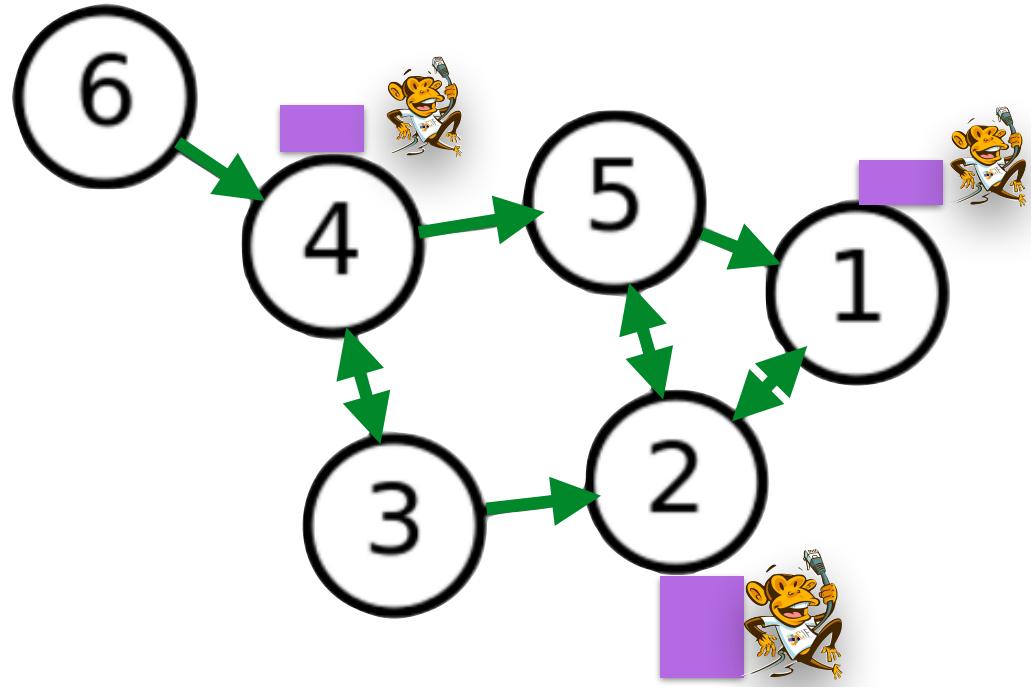
$$x_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad x_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad x_2 = \begin{pmatrix} 0 \\ 0 \\ .5 \\ 0 \\ .5 \\ 0 \end{pmatrix} \quad \rightarrow$$

# 1 step Random Walk

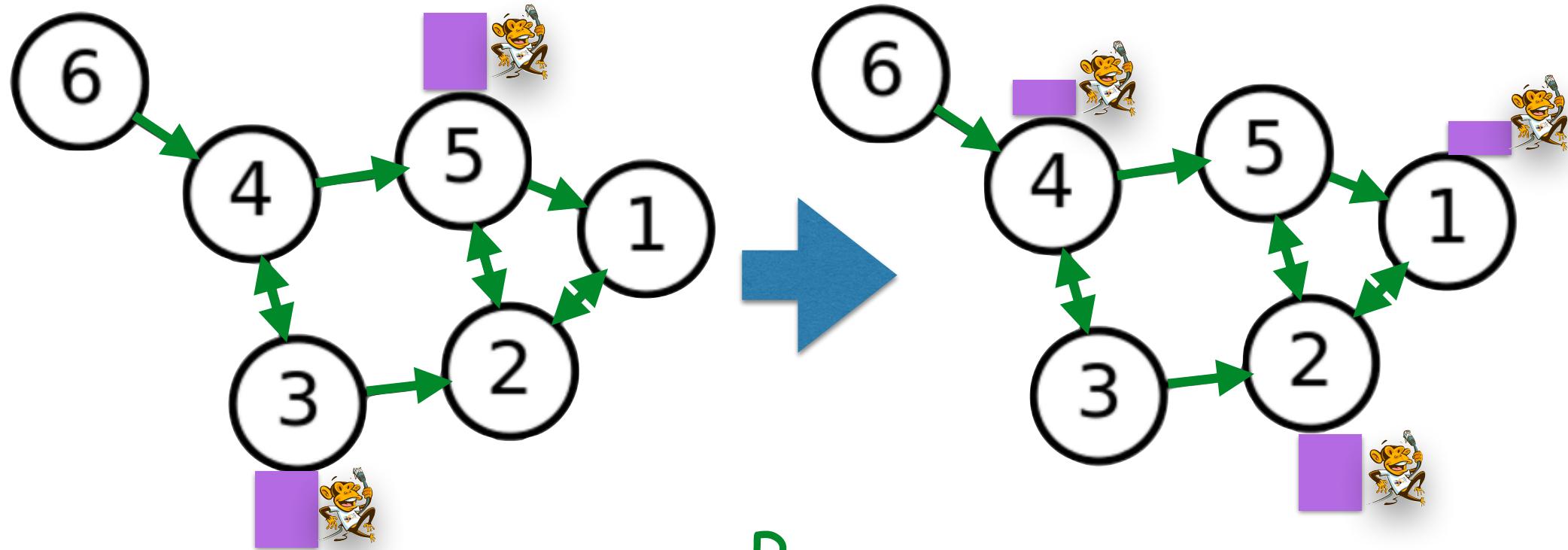


$$x_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad x_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad x_2 = \begin{pmatrix} 0 \\ 0 \\ .5 \\ 0 \\ .5 \\ 0 \end{pmatrix} \quad \rightarrow$$

# 1 step Random Walk



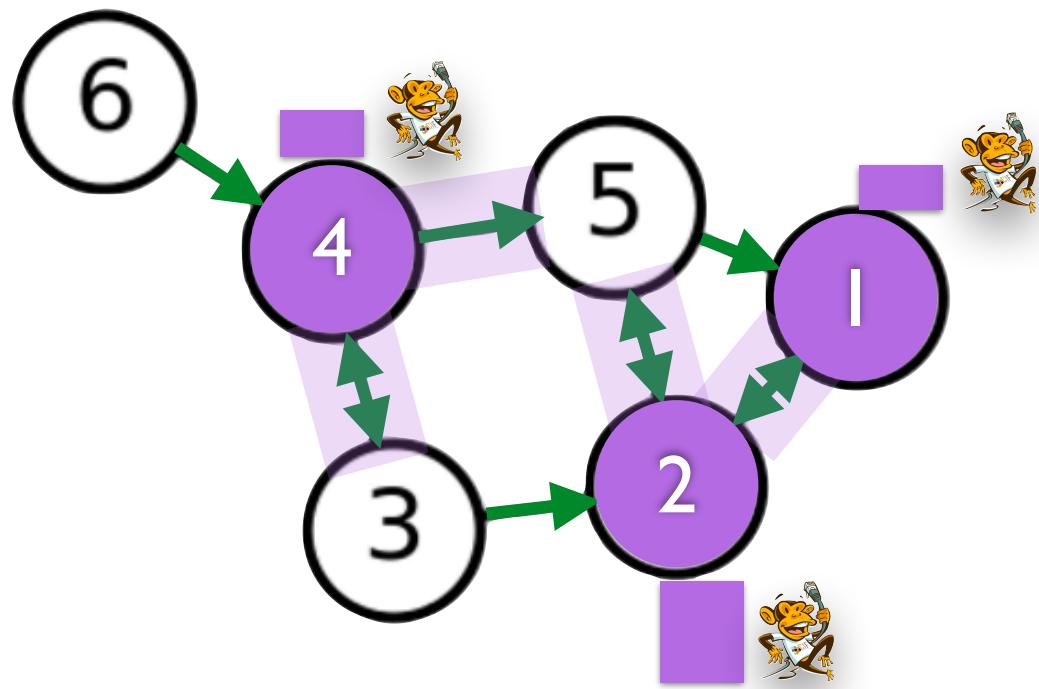
$$x_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad x_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad x_2 = \begin{pmatrix} 0 \\ 0 \\ .5 \\ 0 \\ .5 \\ 0 \end{pmatrix} \quad \rightarrow \quad x_3 = \begin{pmatrix} .25 \\ .5 \\ 0 \\ .25 \\ 0 \\ 0 \end{pmatrix}$$


 $P$ 
 $x_3$ 
 $.25$   
 $.5$   
 $0$   
 $.25$   
 $0$   
 $0$ 
 $=$ 

$$\begin{pmatrix} 0 & .5 & 0 & 0 & .5 & 0 \\ 1 & 0 & .5 & 0 & .5 & 0 \\ 0 & 0 & 0 & .5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & .5 & 0 & .5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

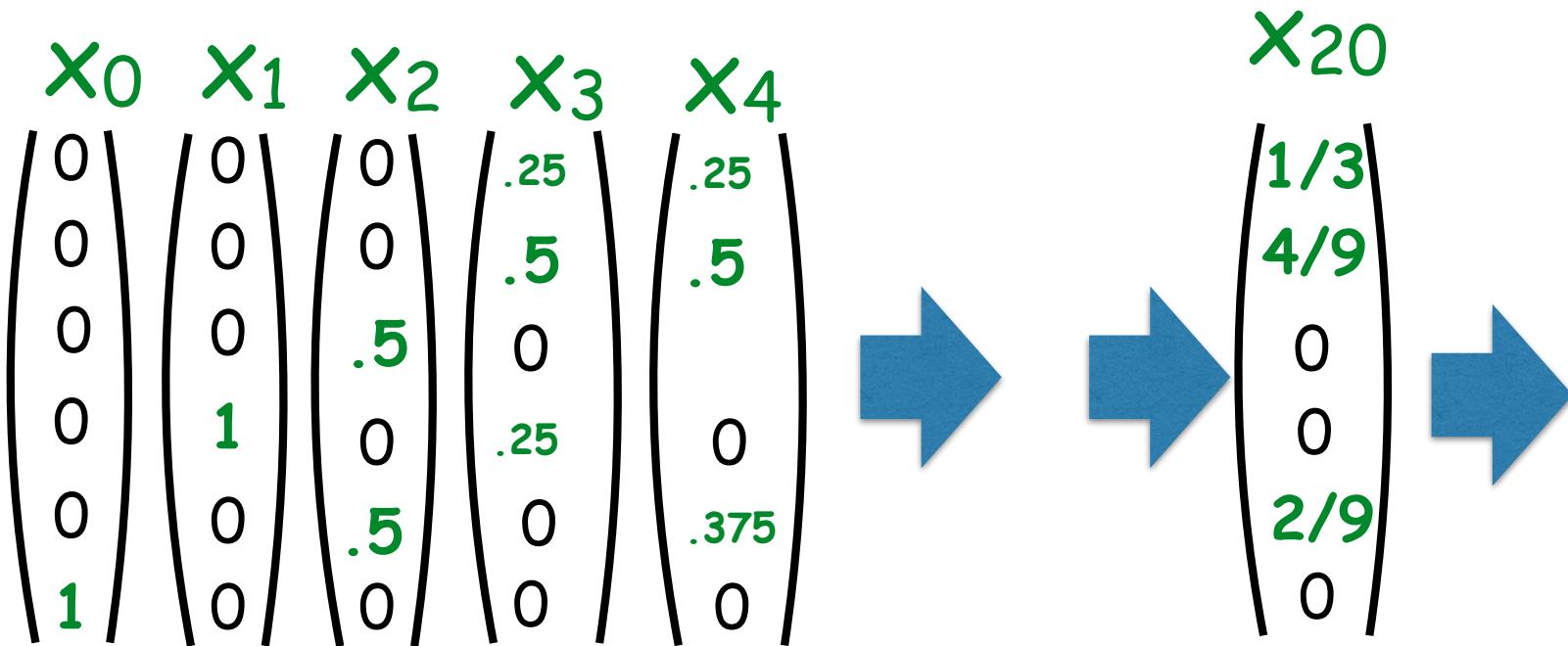
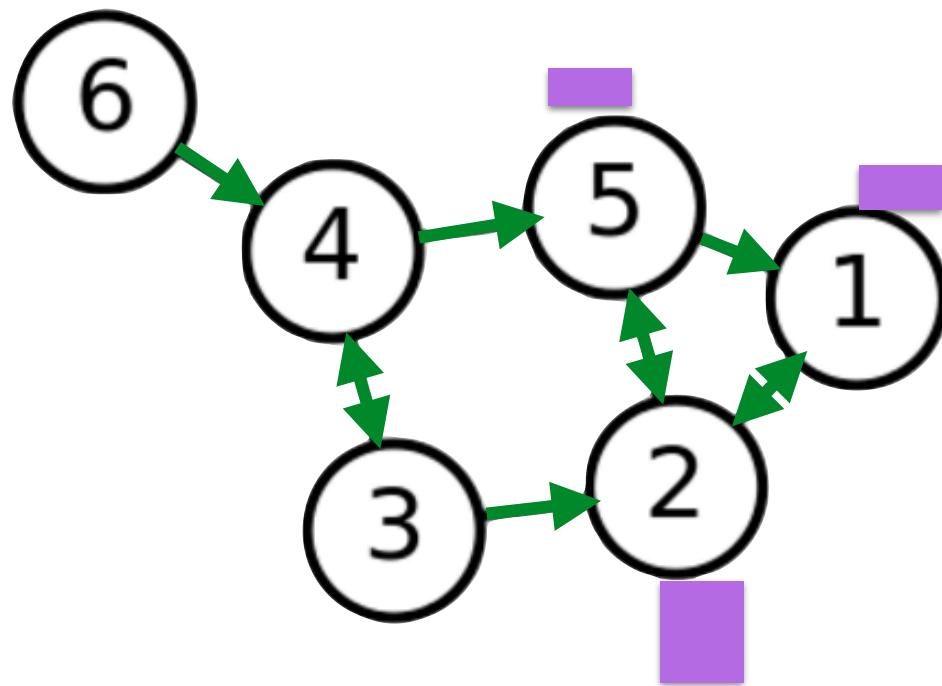
 $x_2$ 
 $0$   
 $0$   
 $.5$   
 $0$   
 $.5$   
 $0$ 
 $x$

# 1 step Random Walk

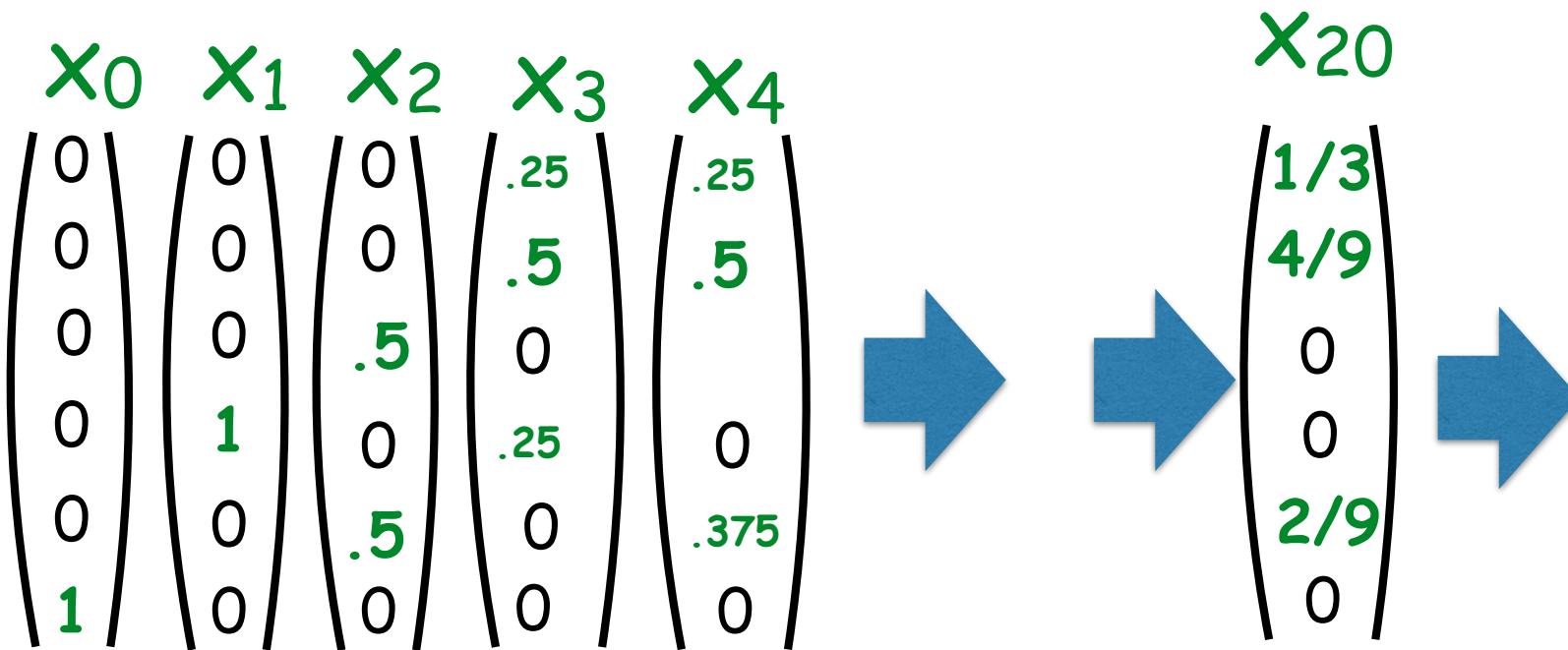
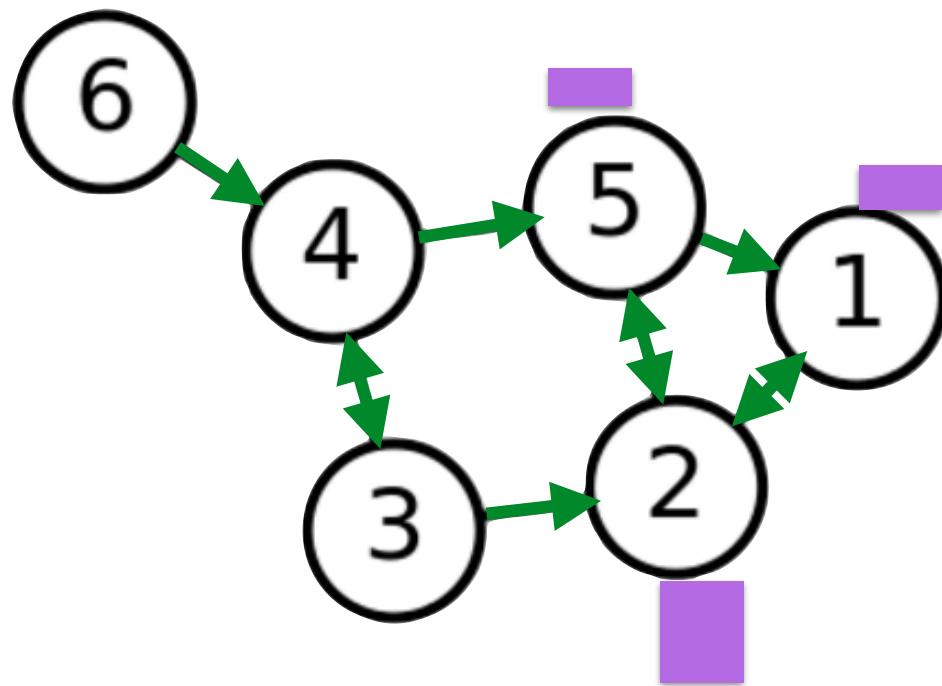


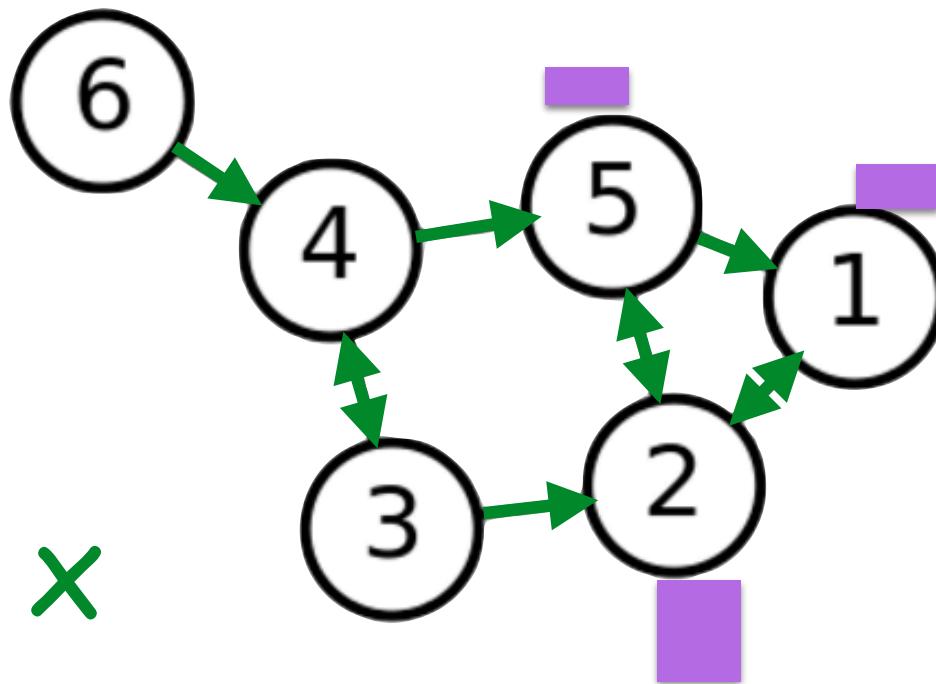
$$x_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad x_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad x_2 = \begin{pmatrix} 0 \\ 0 \\ .5 \\ 0 \\ .5 \\ 0 \end{pmatrix} \quad x_3 = \begin{pmatrix} .25 \\ .5 \\ 0 \\ .25 \\ 0 \\ 0 \end{pmatrix} \quad \xrightarrow{\text{blue arrow}} \quad x_4 = \begin{pmatrix} .25 \\ .5 \\ .125 \\ 0 \\ 0 \\ .375 \end{pmatrix}$$

# 1 step Random Walk



# 1 step Random Walk



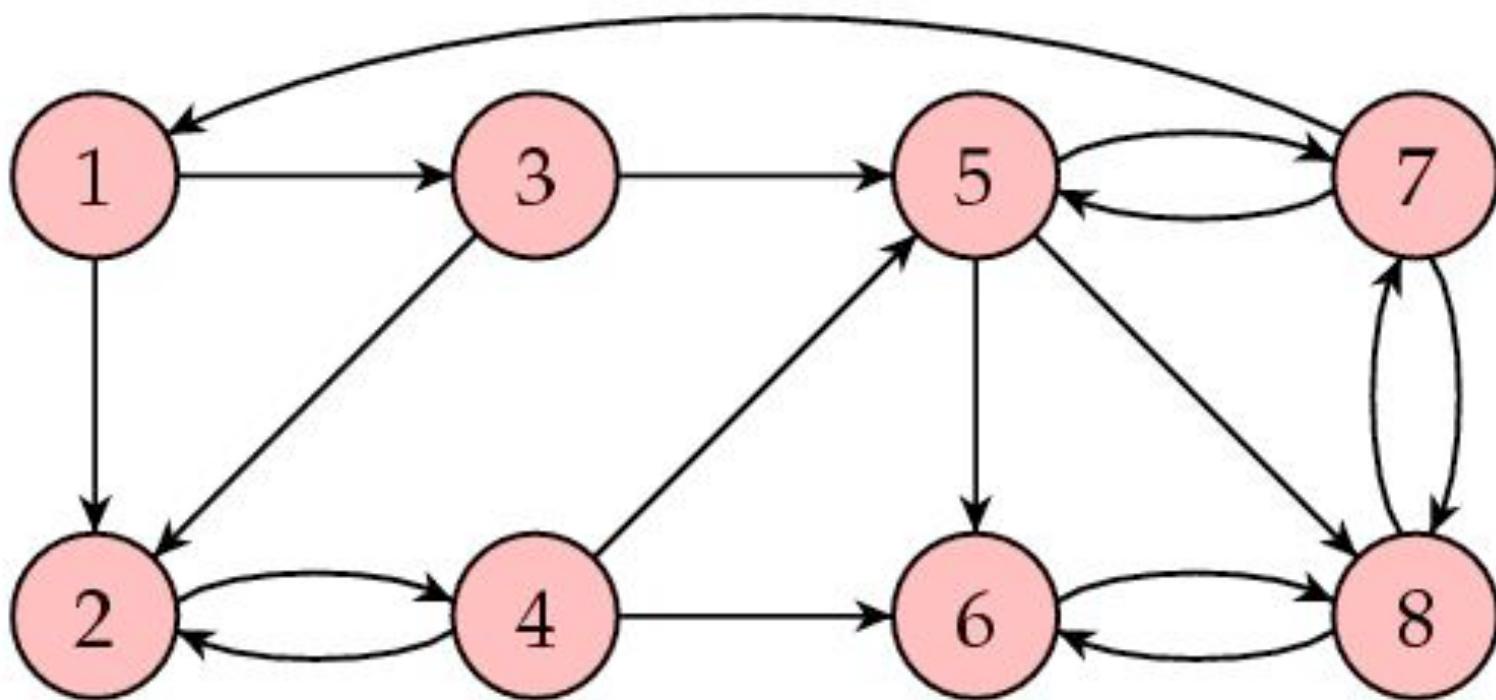


$$P \times = x$$

$$\begin{array}{c}
 P \\
 \left( \begin{array}{cccccc}
 x_{21} & & & & & & x_{20} \\
 \left| \begin{array}{c} 1/3 \\ 4/9 \\ 0 \\ 0 \\ 2/9 \\ 0 \end{array} \right| & = & \left( \begin{array}{cccccc}
 0 & .5 & 0 & 0 & .5 & 0 \\
 1 & 0 & .5 & 0 & .5 & 0 \\
 0 & 0 & 0 & .5 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 1 \\
 0 & 0 & .5 & 0 & 0 & 1 \\
 0 & .5 & 0 & .5 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0
 \end{array} \right) & \left| \begin{array}{c} 1/3 \\ 4/9 \\ 0 \\ 0 \\ 2/9 \\ 0 \end{array} \right| \\
 \end{array} \right)$$

$$P x = \lambda x$$

$\lambda=1$   $x$  is an Eigen vector

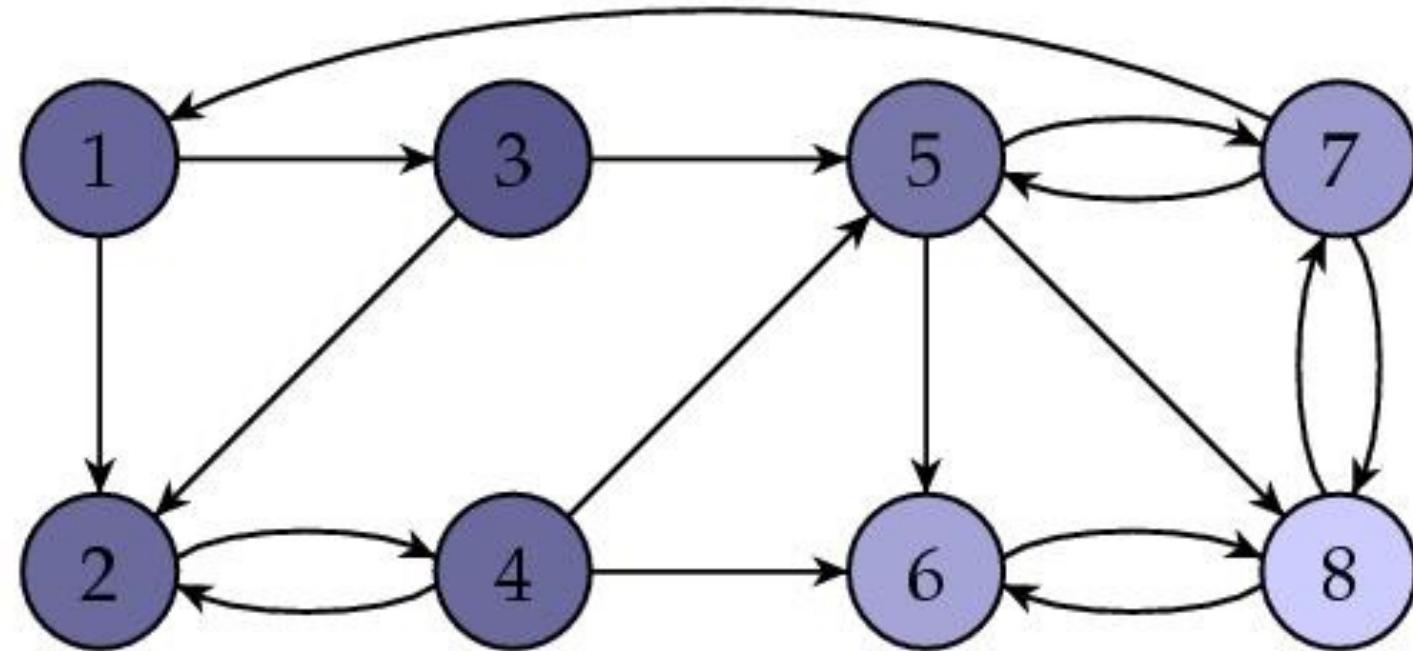
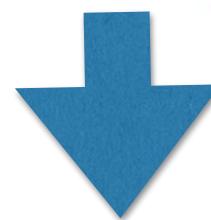
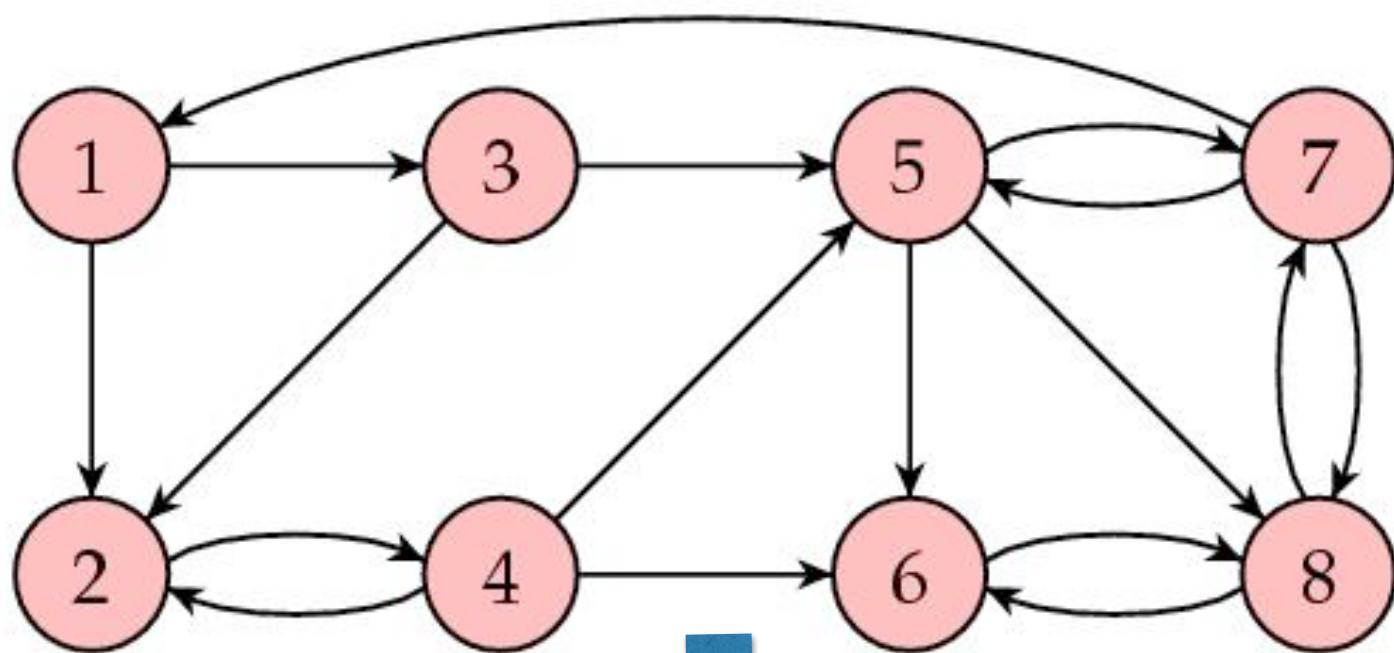


<http://www.ams.org/samplings/feature-column/fcarc-pagerank>

The corresponding matrix is

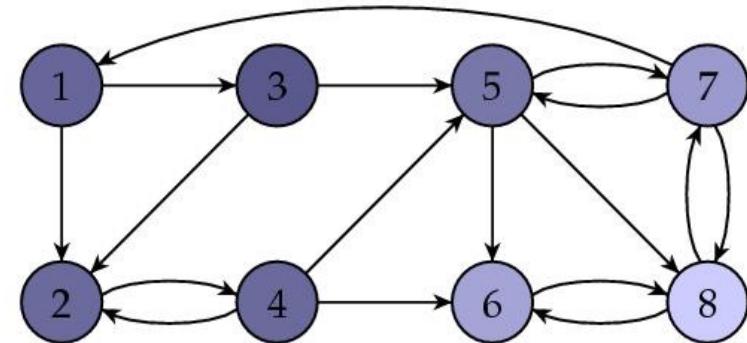
$$\mathbf{P} = \begin{bmatrix}
 0 & 0 & 0 & 0 & 0 & 0 & 1/3 & 0 \\
 1/2 & 0 & 1/2 & 1/3 & 0 & 0 & 0 & 0 \\
 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 1/2 & 1/3 & 0 & 0 & 1/3 & 0 \\
 0 & 0 & 0 & 1/3 & 1/3 & 0 & 0 & 1/2 \\
 0 & 0 & 0 & 0 & 1/3 & 0 & 0 & 1/2 \\
 0 & 0 & 0 & 0 & 1/3 & 1 & 1/3 & 0
 \end{bmatrix}$$

with stationary vector  $\mathbf{x} = \begin{bmatrix} 0.0600 \\ 0.0675 \\ 0.0300 \\ 0.0675 \\ 0.0975 \\ 0.2025 \\ 0.1800 \\ 0.2950 \end{bmatrix}$



# How to Compute $x$

$$x_{k+1} = P x_k$$



$x_0 \ x_1 \ x_2 \ x_3 \ x_4 \ x_{60} \ x_{61}$

|   |     |      |        |        |     |        |        |
|---|-----|------|--------|--------|-----|--------|--------|
| 1 | 0   | 0    | 0      | 0.0278 | ... | 0.06   | 0.06   |
| 0 | 0.5 | 0.25 | 0.1667 | 0.0833 | ... | 0.0675 | 0.0675 |
| 0 | 0.5 | 0    | 0      | 0      | ... | 0.03   | 0.03   |
| 0 | 0   | 0.5  | 0.25   | 0.1667 | ... | 0.0675 | 0.0675 |
| 0 | 0   | 0.25 | 0.1667 | 0.1111 | ... | 0.0975 | 0.0975 |
| 0 | 0   | 0    | 0.25   | 0.1806 | ... | 0.2025 | 0.2025 |
| 0 | 0   | 0    | 0.0833 | 0.0972 | ... | 0.18   | 0.18   |
| 0 | 0   | 0    | 0.0833 | 0.3333 | ... | 0.295  | 0.295  |

# Sink Node Problem



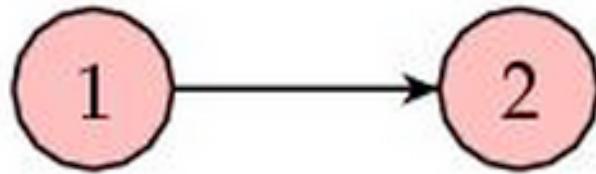
with matrix  $P = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$

Here is one way in which our algorithm could proceed:

| $x_0$ | $x_1$ | $x_2$ | $x$ |
|-------|-------|-------|-----|
| 1     | 0     | 0     | 0   |
| 0     | 1     | 0     | 0   |

why? how to solve?

# Solution



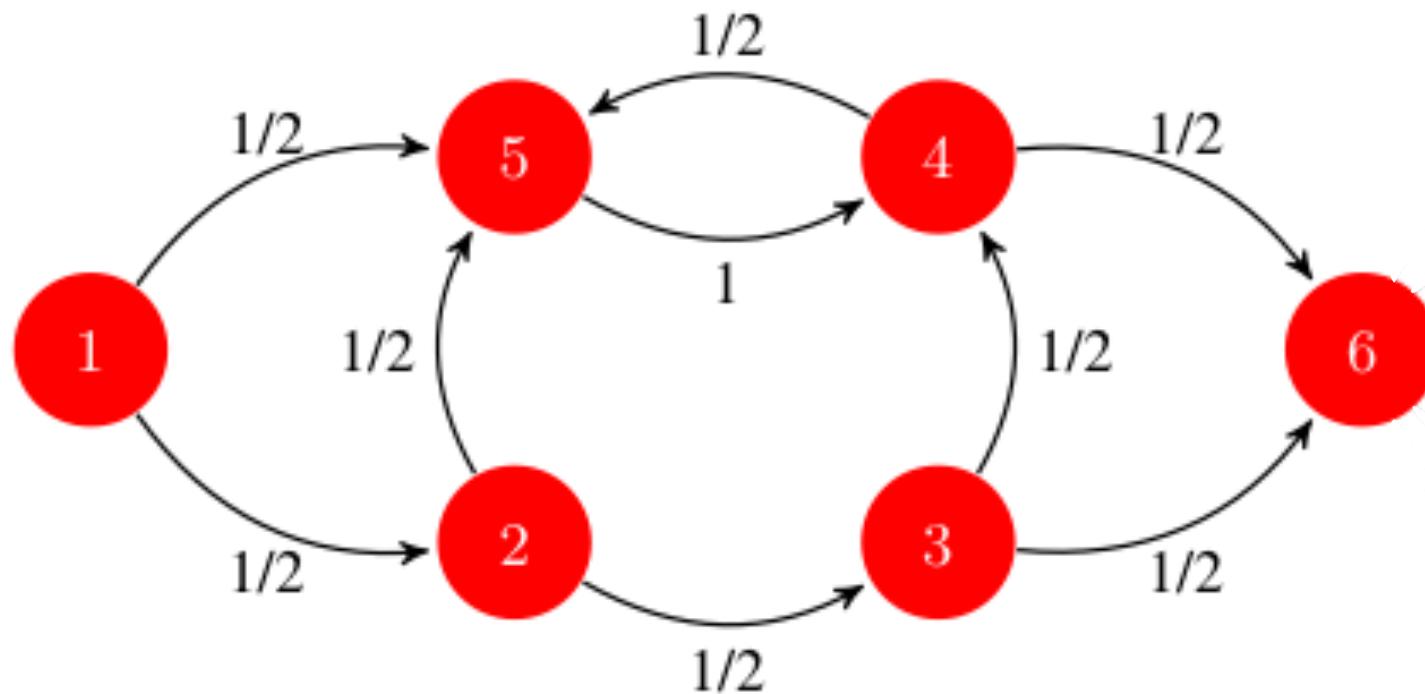
with matrix

$$P = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$



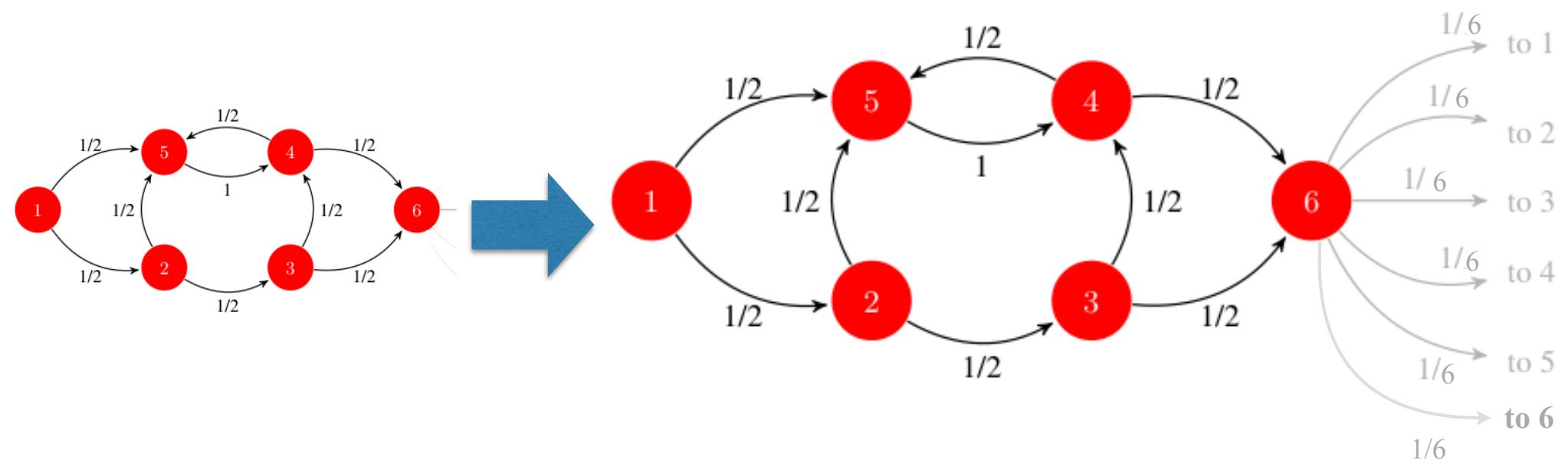
with matrix  $S = \begin{bmatrix} 0 & 1/2 \\ 1 & 1/2 \end{bmatrix}$  and eigenvector  $X = \begin{bmatrix} 1/3 \\ 2/3 \end{bmatrix}$

# Sink Node Problem



matrix P

# Solving Sink Node Problem



matrix  $P$



matrix  $S$

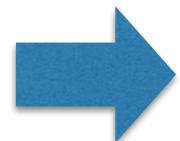
# Solving Sink Node Problem

sink node

$$\left| \begin{array}{cccccc} 0 & .5 & 0 & 0 & .5 & 0 \\ 1 & 0 & .5 & 0 & .5 & 0 \\ 0 & 0 & 0 & .5 & 0 & 0 \\ 0 & 0 & .5 & 0 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right|$$

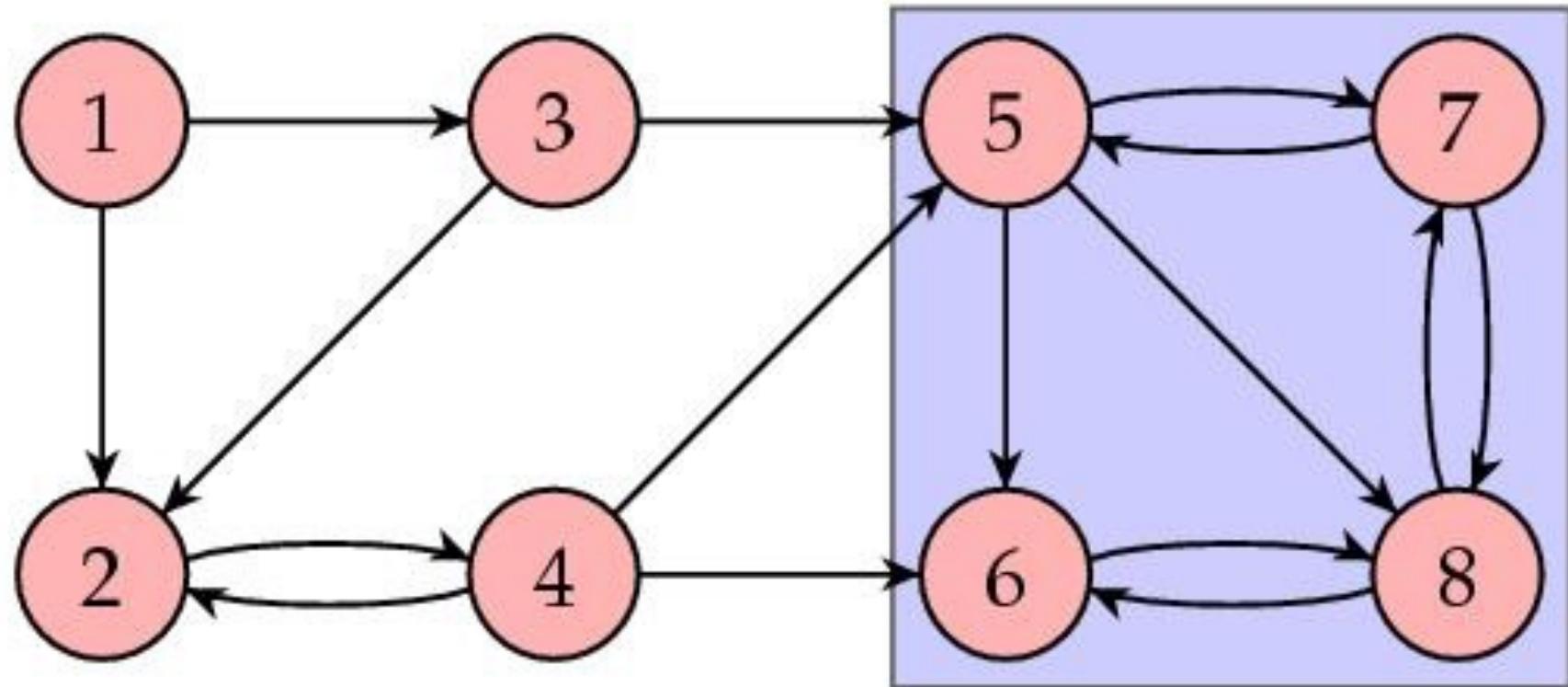
$$\left| \begin{array}{cccccc} 0 & .5 & 0 & 0 & .5 & 1/6 \\ 1 & 0 & .5 & 0 & .5 & 1/6 \\ 0 & 0 & 0 & .5 & 0 & 1/6 \\ 0 & 0 & .5 & 0 & 0 & 1/6 \\ 0.5 & 0.5 & 0 & 0 & 0 & 1/6 \\ 0 & 0 & 0 & 0 & 0 & 1/6 \end{array} \right|$$

matrix P

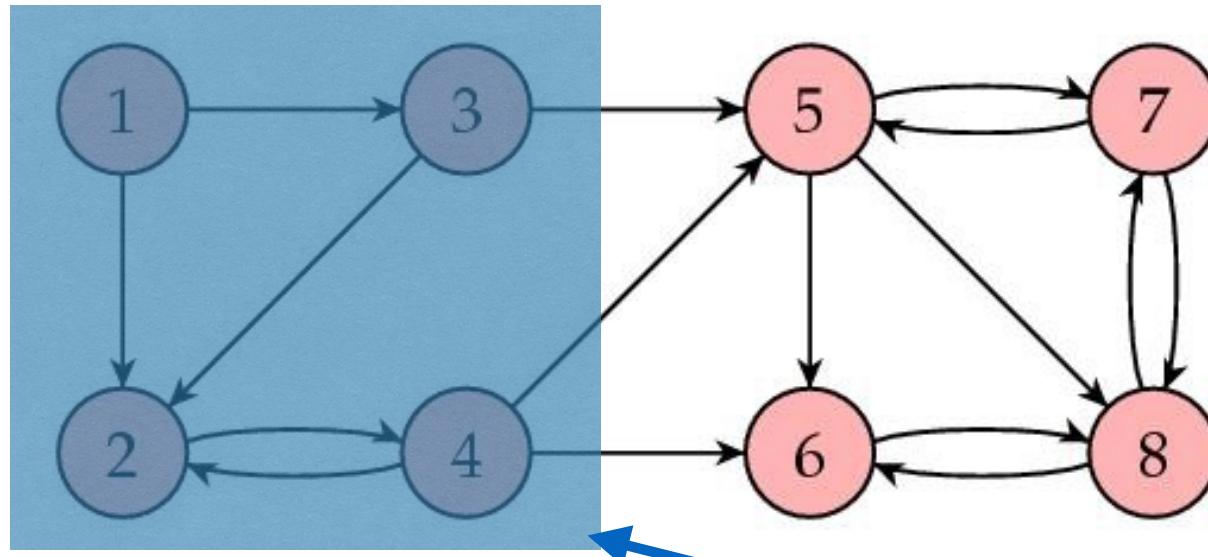


matrix S

# Sink Region Problem



# Sink Region Problem



$$S = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 1/3 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 1/3 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1/3 & 1/3 & 0 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 1/3 & 0 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 1/3 & 1 & 1/2 & 0 \end{bmatrix} \quad X = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0.12 \\ 0.24 \\ 0.24 \\ 0.4 \end{bmatrix}$$

$\times$

# Solving Sink Region Problem

"randomly follow a link"

"randomly jump to a node"

$$\begin{pmatrix} 0 & .5 & 0 & 0 & .5 & 1 \\ 1 & 0 & .5 & 0 & .5 & 0 \\ 0 & 0 & 0 & .5 & 0 & 0 \\ 0 & 0 & .5 & 0 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

matrix  $S$

$$\begin{pmatrix} 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & \cdots & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 \end{pmatrix}$$

matrix  $\frac{1}{n}\mathbf{1}$

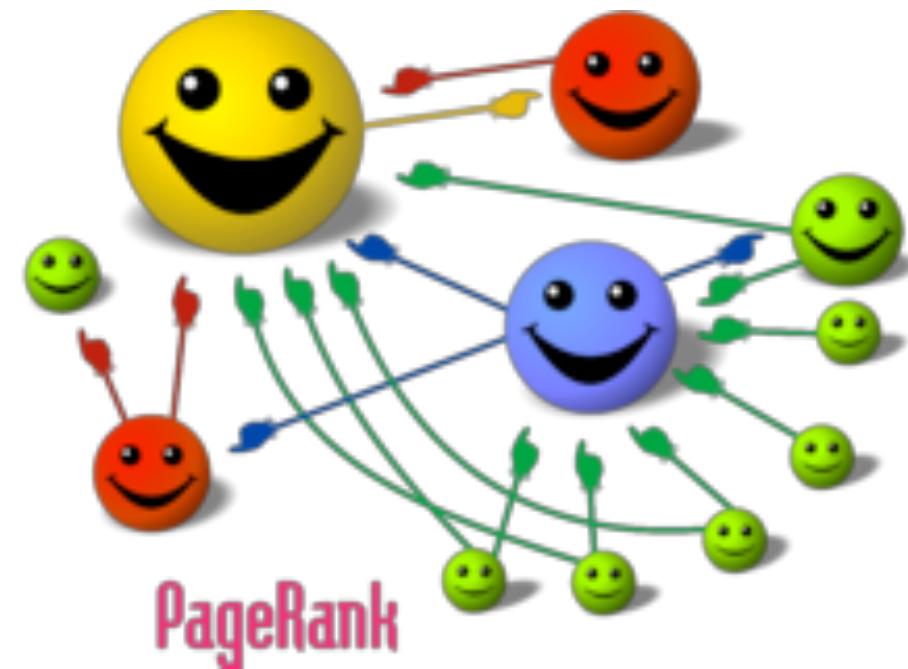
$n \times n$ ,  
all-one  
matrix

# Final Solution

Adding outlinks for every node

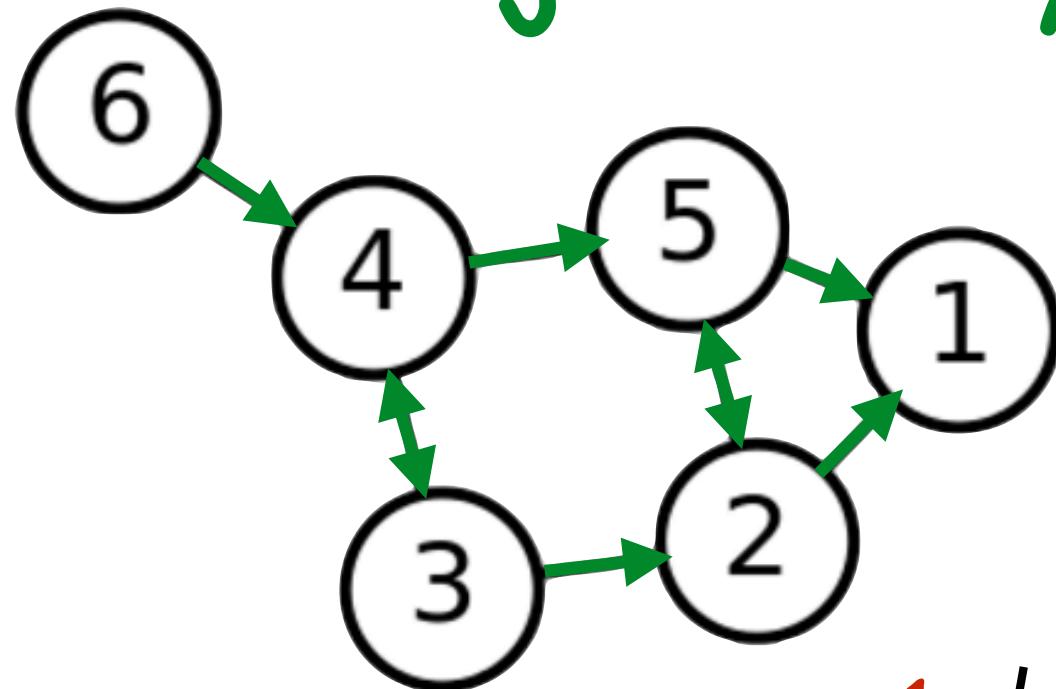
$$G = a S + (1-a) \frac{1}{n} 1$$

a is a number between 0 and 1



a: probability of  
randomly following  
a link

# Adjacency Matrix (A)

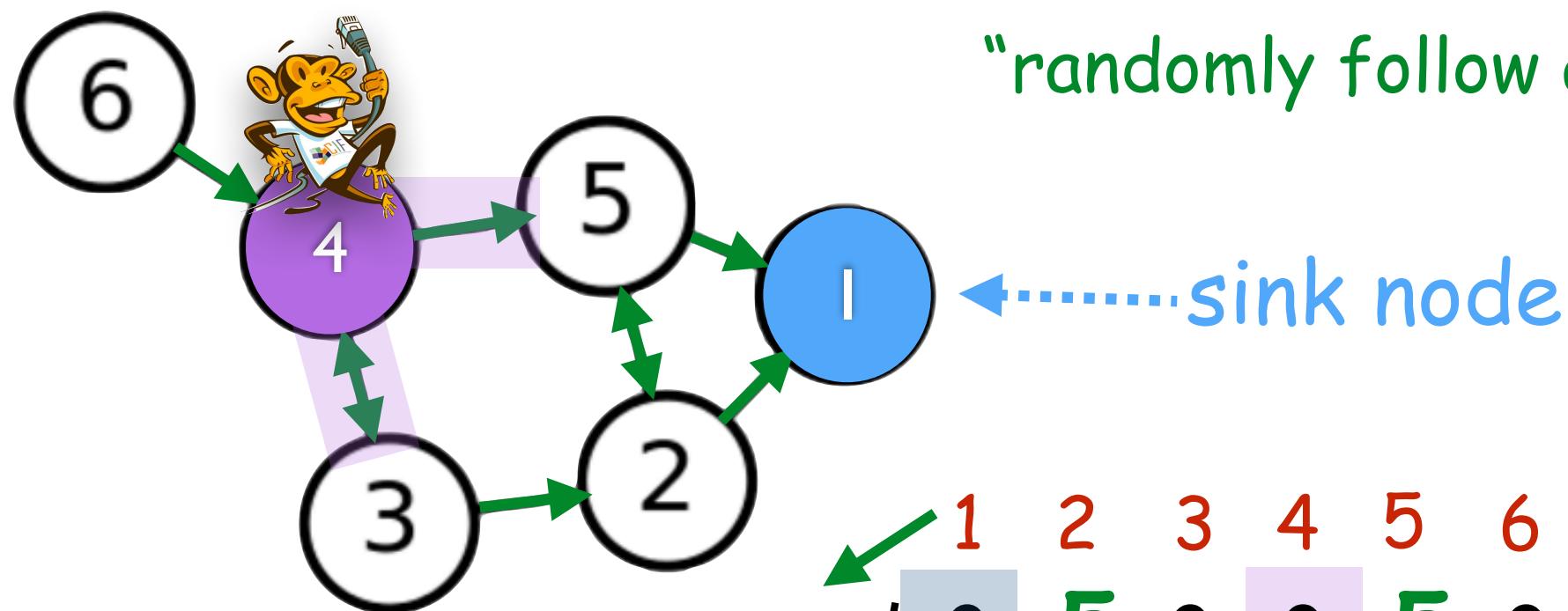


$A =$

|   | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | 0 | 1 | 0 | 0 | 1 | 0 |
| 2 | 0 | 0 | 1 | 0 | 1 | 0 |
| 3 | 0 | 0 | 0 | 1 | 0 | 0 |
| 4 | 0 | 0 | 1 | 0 | 0 | 0 |
| 5 | 0 | 1 | 0 | 1 | 0 | 1 |
| 6 | 0 | 0 | 0 | 0 | 0 | 0 |

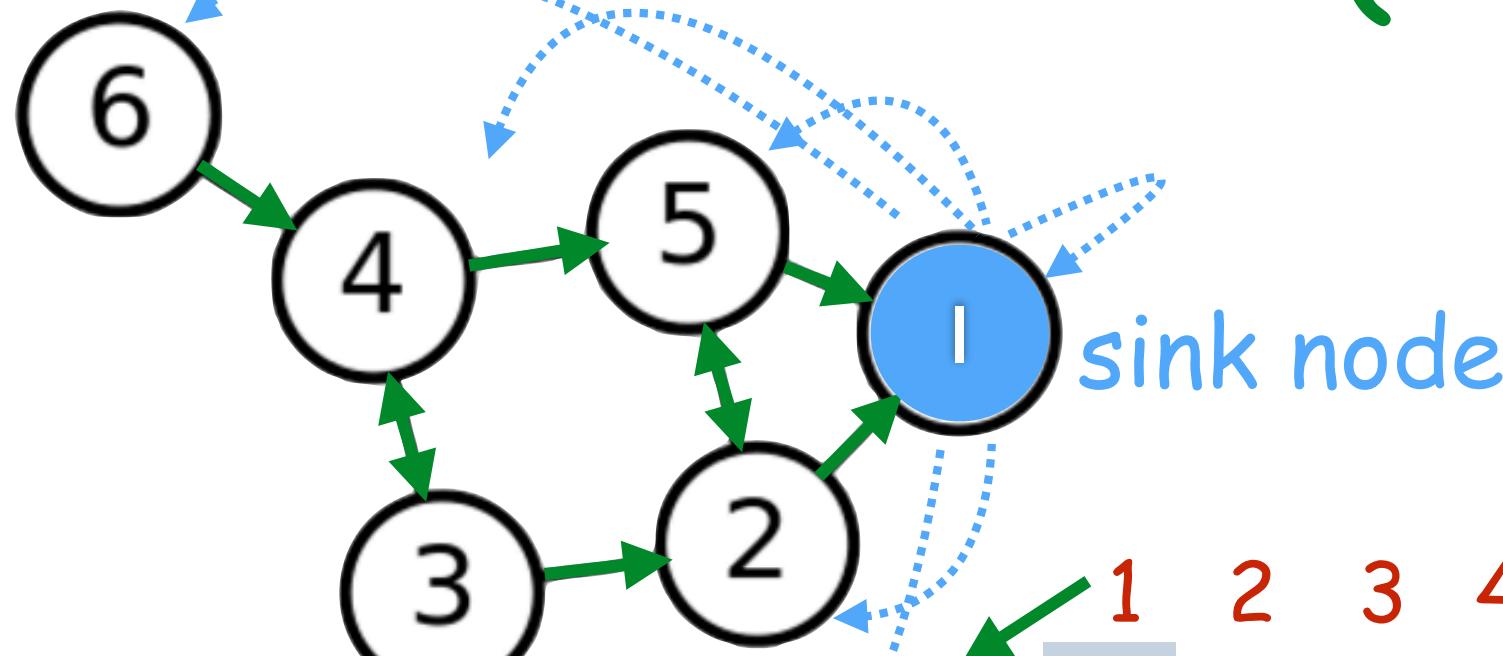
# Transition Matrix ( $P$ )

"randomly follow a link"



$$P = \begin{pmatrix} & 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 0 & .5 & 0 & 0 & .5 & 0 \\ 2 & 0 & 0 & .5 & 0 & .5 & 0 \\ 3 & 0 & 0 & 0 & .5 & 0 & 0 \\ 4 & 0 & 0 & .5 & 0 & 0 & 0 \\ 5 & 0 & .5 & 0 & .5 & 0 & 1 \\ 6 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

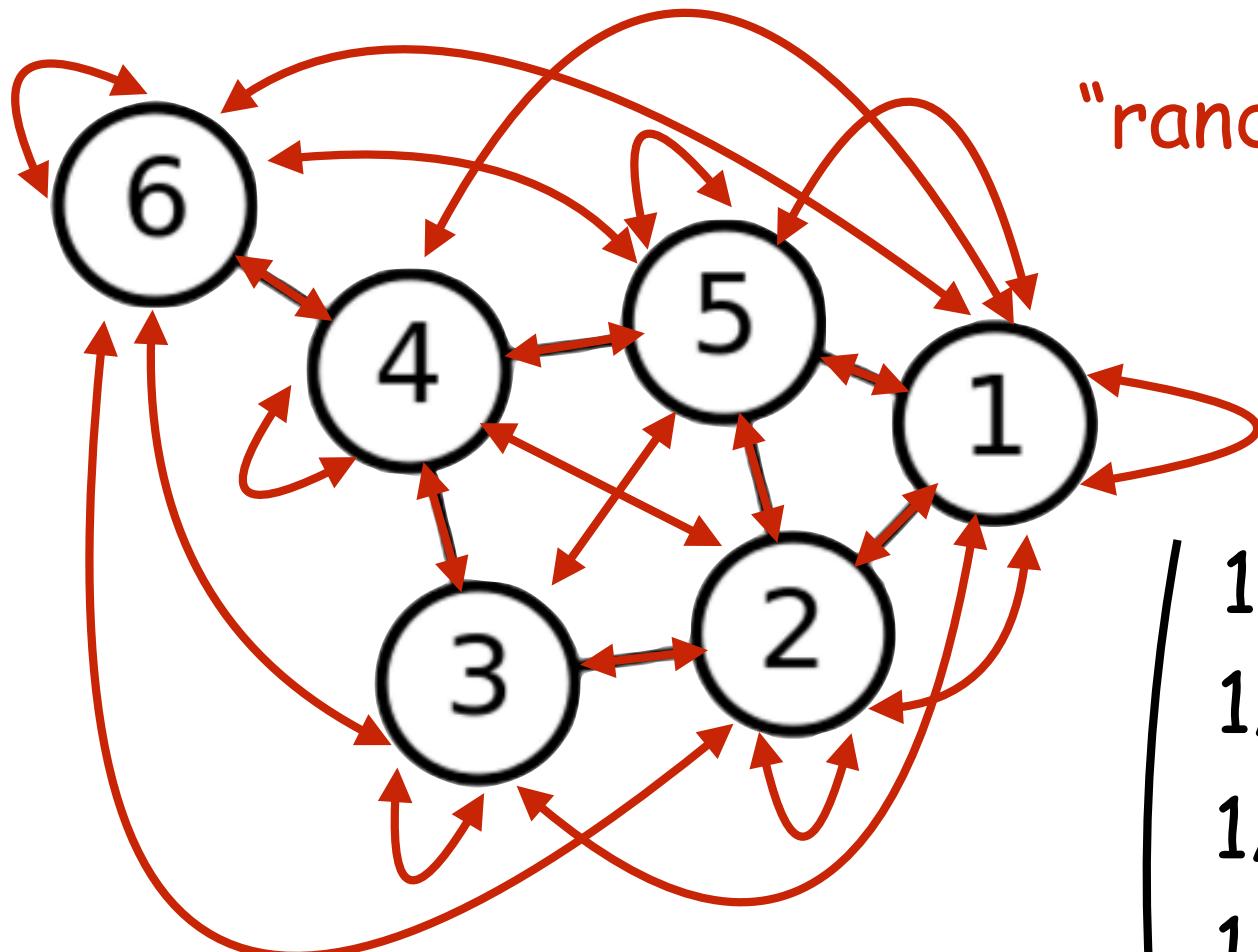
# Matrix ( $S$ )



$S =$

| 1 | 2   | 3  | 4  | 5  | 6  |   |
|---|-----|----|----|----|----|---|
| 1 | 1/6 | .5 | 0  | 0  | .5 | 0 |
| 2 | 1/6 | 0  | .5 | 0  | .5 | 0 |
| 3 | 1/6 | 0  | 0  | .5 | 0  | 0 |
| 4 | 1/6 | 0  | .5 | 0  | 0  | 0 |
| 5 | 1/6 | .5 | 0  | .5 | 0  | 1 |
| 6 | 1/6 | 0  | 0  | 0  | 0  | 0 |

# Matrix ( $\frac{1}{n} \mathbf{1}$ )



"randomly jump to a node"

$$\begin{pmatrix} 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 \\ 1/6 & \cdots & 1/6 \\ 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 \end{pmatrix}$$

$$\text{Matrix } G = a S + (1-a) \frac{1}{n} \mathbf{1}$$

"randomly follow a link"

$$.8 \times \begin{pmatrix} 1/6 & .5 & 0 & 0 & .5 & 1 \\ 1/6 & 0 & .5 & 0 & .5 & 0 \\ 1/6 & 0 & 0 & .5 & 0 & 0 \\ 1/6 & 0 & .5 & 0 & 0 & 0 \\ 1/6 & .5 & 0 & 0 & 0 & 0 \\ 1/6 & 0 & .5 & 0 & 0 & 0 \end{pmatrix}$$

"randomly jump to a node"

$$+.2 \times \begin{pmatrix} 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & \cdots & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 \end{pmatrix}$$

$$\text{Matrix } G = a S + (1-a) \frac{1}{n} \mathbf{1}$$

