

# Wrap Up and PageRank

CSCI 10 - Santa Clara University - Fall 2016  
Michael J. Bannister

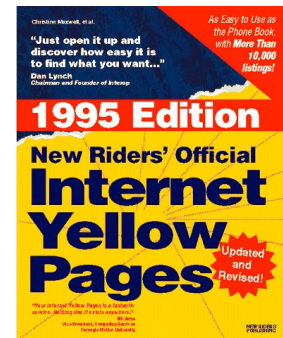
## Announcements

- Report due today (anytime)
- Homework due now
- Time slots for talks will be 4 to 8 minutes.

## Google's PageRank

## The Problem

Assuming we have already collected the contents of a large collection of webpages, how do we help users find sites of interest to them?



# Keyword Solution

When the user performs a search find those site containing each word in the search string and sort site by how often the keywords appear.

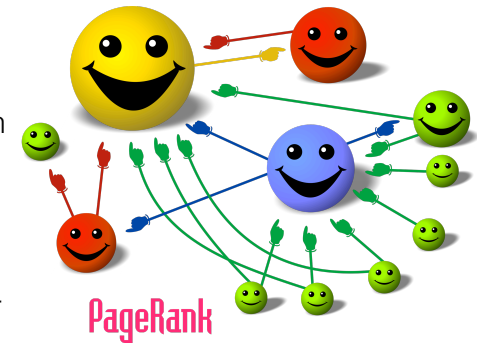


This led to abuse quickly: site started putting fake keyword hidden on their web sites rendering search results useless.

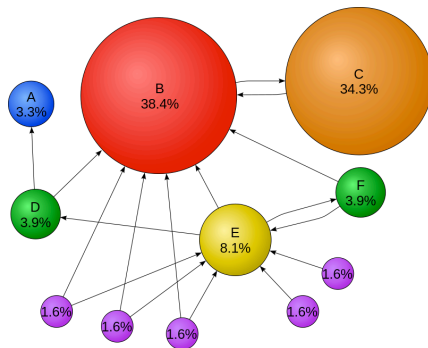
# Page Rank

In the late 1990s Larry Page and Sergey Brin found a method of combining keyword search with a popularity metric.

So when you search for a site you get the most popular site matching your keywords.



# Random Surfer Model



What is the probability that a random web surfer landed on a given webpage.

# Impact on Society



Essentially all known information is now easily accessible!

Is there any reason left to wonder?

# Societies Impact

People have started to create fake webpages to boost PageRank and in turn Google is combating this.



A never ending battle between search and scammers!

Thank You!

Any questions?

Continue with Grade  
Example