

Social Networks

Luigi Giugliano¹, Steven Rosario Sirchia¹

¹Università degli studi di Salerno

6 giugno 2016

Introduction

The purpose of our work is to test different algorithms in the three fundamental areas for assembling any search engine offering a Sponsored Search system:

- **Ranking** of web documents
- **Matching** of words inside documents
- **Auctions** for acquiring advertisement slots.

We will briefly talk about the proposed algorithms, and then compare running times and results obtained from their execution more in detail, suggesting what combination of algorithms seems to be the best for realizing a new search engine.

Overview

1 Ranking

- Page Rank
- HITS
- Results

2 Matching

- Best Match
- Improved Best Match
- Results

3 Search Engine

4 Auction

- First Price Auction
- Generalized Second Price Auction
- Results

Overview

1 Ranking

- Page Rank
- HITS
- Results

2 Matching

- Best Match
- Improved Best Match
- Results

3 Search Engine

4 Auction

- First Price Auction
- Generalized Second Price Auction
- Results

Page Rank

Page Rank

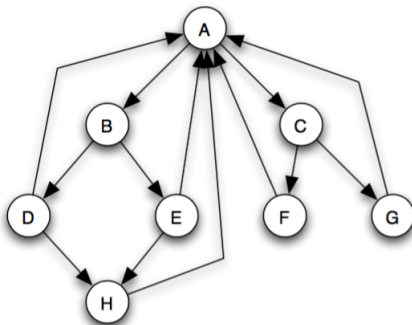
The intuition behind *Page Rank* is:

“a page is important if it is cited by other important pages”.

This intuition rises from the usual endorsement mode, for example, among academic or governmental pages, among bloggers, or among personal pages more generally. It is also the dominant mode in the scientific literature.

Algorithm

We can think of PageRank as a kind of “fluid” that circulates through the network, passing from node to node across edges, and pooling at the nodes that are the most important.



Algorithm Steps

- In a network with n nodes, we assign all nodes the same initial PageRank, set to be $1/n$.

Algorithm Steps

- In a network with n nodes, we assign all nodes the same initial PageRank, set to be $1/n$.
- We choose a number of steps k .

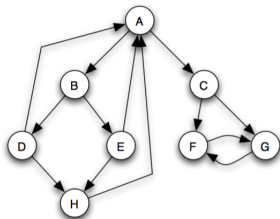
Algorithm Steps

- In a network with n nodes, we assign all nodes the same initial PageRank, set to be $1/n$.
- We choose a number of steps k .
- We then perform a sequence of k updates to the PageRank values, using the following rule for each update:

Basic PageRank Update Rule: Each page divides its current PageRank equally across its out-going links, and passes these equal shares to the pages it points to. (If a page has no out-going links, it passes all its current PageRank to itself.) Each page updates its new PageRank to be the sum of the shares it receives.

The “Wrong” nodes

There is a difficulty with the basic definition of PageRank, however: in many networks, the “wrong” nodes can end up with all the PageRank.



The Wrong nodes are a small sets of nodes that can be reached from the rest of the graph, but have no paths back.

Scaled PageRank

We can use the mechanism of fluid presented above, there is a **counter-balancing process** preventing that all the water stands only on downhill places on the earth.

Scaled PageRank Update Rule

First apply the Basic PageRank Update Rule.

Then scale down all PageRank values by a factor of s , shrinking the total from 1 to s .

We divide the residual $1 - s$ units of PageRank equally over all nodes, giving $(1 - s)/n$ to each.

Results

We are going to present the result of the experiment, comparing the **execution time** of PageRank on following inputs:

- Graph of 1000 nodes
- Graph of 2000 nodes
- Graph of 5000 nodes
- Graph of 10000 nodes
- Graph of 20000 nodes
- Full Graph (30000 nodes)

All the graphs are generated by chunking the Full Graph.

Results



Overview

1 Ranking

- Page Rank
- HITS
- Results

2 Matching

- Best Match
- Improved Best Match
- Results

3 Search Engine

4 Auction

- First Price Auction
- Generalized Second Price Auction
- Results

Overview

- 1 Ranking
 - Page Rank
 - HITS
 - Results
- 2 Matching
 - Best Match
 - Improved Best Match
 - Results
- 3 Search Engine
- 4 Auction
 - First Price Auction
 - Generalized Second Price Auction
 - Results

Overview

- 1 Ranking
 - Page Rank
 - HITS
 - Results
- 2 Matching
 - Best Match
 - Improved Best Match
 - Results
- 3 Search Engine
- 4 Auction
 - First Price Auction
 - Generalized Second Price Auction
 - Results

