

Page Rank

Mason U'Ren, James Musselman, Zander Nelson

October 10, 2016

Abstract

A page rank results from a mathematical algorithm based on the web-graph, created by all Internet sites as nodes and hyper-links as edges. This rank value indicates an importance of a particular page where a hyper-link to a page counts as a vote of support. The page rank is then defined recursively and depends on the number and page rank metric of all pages that link to it - "incoming links." ie. a page that is linked to by many pages with high page rank receives a high rank itself.

Algorithm 1 PageRank Algorithm

```
1: procedure PAGERANK( $G$ ,  $iteration$ )
2:    $d \leftarrow 0.85$                                  $\triangleright$  G: in-link file, iteration: number of iteration
3:    $oh \leftarrow G$                                  $\triangleright$  damping factor: 0.85
4:    $ih \leftarrow G$                                  $\triangleright$  get in-link hash from G
5:    $N \leftarrow G$                                  $\triangleright$  get number of pages from G
6:   for all  $p$  in the graph do
7:      $opg[p] \leftarrow \frac{1}{N}$                                  $\triangleright$  initialize PageRank
8:   end for
9:   while  $iteration > 0$  do
10:     $dp \leftarrow 0$ 
11:    for all  $p$  that has no out-links do
12:       $dp \leftarrow dp + d * \frac{opg[p]}{N}$   $\triangleright$  get PageRank from pages without out-links
13:    end for
14:    for all  $p$  in the graph do
15:       $npg[p] \leftarrow dp + \frac{1-d}{N}$                                  $\triangleright$  get PageRank from random jump
16:      for all  $ip$  in  $ih[p]$  do
17:         $npg[p] \leftarrow npg[p] + \frac{d * opg[ip]}{oh[ip]}$                                  $\triangleright$  get PageRank from in-links
18:      end for
19:    end for
20:     $opg \leftarrow npg$                                  $\triangleright$  update PageRank
21:     $iteration \leftarrow iteration - 1$ 
22:  end while
23: end procedure
```

1 Does the algorithm always terminate?

Page rank will stop iterating when convergence is reached. The algorithm actively checks for and stops on convergence by keeping track of the difference in value between each iteration. After many iterations the algorithm reaches a "steady state" where the page rank of all the pages in the indexed converged. Once this happens the algorithm terminates.