The PageRank Citation Ranking
REFERENCES


The PageRank citation ranking: bringing order to the Web.

Introduction

- Traditional web search engines feedback base on key words match:
  - subjective
  - expensive to build and maintain
  - slow to improve
  - cannot cover all esoteric topics

- Simple PageRank Algorithm
Architecture of Link Analysis Search Engine

Online Part

Query

User Interface

Caching

Indexing and Ranking

Inverted Index

Cached Pages

Page & Site Statistics

Index Builder

Web Page Parser

Crawler

Web Graph Builder

Page Weights

Link Analysis

Web Graph

Offline Part
The Simplified PageRank Algorithm Example (1)

Assume a small universe of four web pages: A, B, C and D.

All PageRank $PR(\ )$ in this simplistic system would thus gather to A because all links would be pointing to A.

PageRank (A) = PageRank (B) + PageRank (C) + PageRank (D)
The value of the link-votes is divided among all the outbound links on a page.

Thus, page B gives a vote worth 0.125 to page A and a vote worth 0.125 to page C. Only one third of D's PageRank is counted for A's PageRank (approximately 0.083).

\[
PR(A) = \frac{PR(B)}{2} + \frac{PR(C)}{1} + \frac{PR(D)}{3}
\]
The PageRank is equal to the document's own PageRank score divided by the number of outbound links.

\[ L() : \text{the number of outbound links from current document.} \]

\[ PR(A) = \frac{PR(B)}{L(B)} + \frac{PR(C)}{L(C)} + \frac{PR(D)}{L(D)} \]
Simplified PageRank-Algorithm

- PageRank value is calculated by web-page
- Web pages hand their rank down to the pages they link to.

\[ PR(u) = \sum_{v \in B_u} \frac{PR(v)}{L(v)} \]

- Predecessor nodes \( B_u \)
- Successor nodes \( F_u \)
- The PageRank value for a page \( u \) is dependent on the PageRank values for each page \( v \) out of the set \( B_u \) (this set contains all pages linking to page \( u \)), divided by the number \( L(v) \) of links from page \( v \)
Search Result Compare (PageRank)

➢ The feedback result is sort by PageRank

  – Sometimes the result isn’t related with query key words closely

➢ PageRank take rank value (web page weight) as average distribute

  – The importance of every link is the same

➢ PageRank: follow like a sheep

  – The result maybe isn’t the best, but it never be the worst
Thanks you