

PageRank vs. Web Traffic

- Important component of PageRank calculation is E
 - A vector over the web pages (used as source of rank)
 - Powerful parameter to adjust the page ranks
- The vector E corresponds to the distribution of web pages that a random surfer periodically jumps to from the search engine
- Some highly accessed web pages have low page rank possibly because
 - People do not want to link to these pages from their own web pages (the example in 1998 PageRank paper is pornographic sites...)
 - Some important backlinks are omitted
 - Use web usage data as a start vector for PageRank.