

What is PageRank?

When learning about Canisius' library database and methods for finding materials, we loosely defined PageRank as an algorithm used by Google to rate certain web pages. The rating would then be used to determine the order in which to display search results. A much more precise definition from wikipedia is as follows:

“PageRank is a link analysis algorithm, named after Larry Page and used by the Google Internet search engine, that assigns a numerical weighting to each element of a hyperlinked set of documents, such as the World Wide Web, with the purpose of 'measuring' its relative importance within the set. The algorithm may be applied to any collection of entities with reciprocal quotations and references. The numerical weight that it assigns to any given element E is referred to as the PageRank of E and denoted by PR(E).” (1)

A Brief History

First and foremost, the PageRank algorithm is not patented by Google. Google has exclusive rights to it, but the patent belongs to Stanford University where it was developed by Larry Page and Sergey Brinn.(1)

The theory behind PageRank is not unique; many similar algorithms were being developed around the same time.

The Algorithm

PageRank is a probability distribution that simply measures the relatively likelihood of a person to click on a link to the desired page. If page Q has a total of three links and one of them goes to P, then the chances of being directed from Q to P is 1/3 (assuming there is an equal chance of clicking any of the three links). For each page that links to P, a simplified rank of P is the summation of all probabilities of all pages that link to P:

$$\text{PageRank}(P) = \sum_{Q \in B_P} \frac{r(Q)}{|Q|}$$

where  $r(Q)$  is the number of links on some  $Q$  that point to  $P$ ,  $B_p$  is the set of all pages pointing to  $P$ , and  $|Q|$  is the number of outlinks from Q.(2)

Conclusion of Pre-Talk

Short of some various details, the algorithm above drives one of the most utilized search engines around the world; yet, it is so stunningly straightforward (for those of us interested in mathematics). In general, this application of link analysis is but a subset of a much larger branch of mathematics; namely, network analysis. To those who would have declared the study of group relations as a waste of time, the patent for the PageRank algorithm was sold for 1.8 million shares of Google; which were sold in 2005 for \$336 million.(3)

References

- (1) "Google Press Center: Fun Facts". [www.google.com](http://www.google.com). Archived from the original on 2009-04-24.
- (2) <http://www.math.umass.edu/~law/Research/PageRank/Google.pdf>
- (3) Lisa M. Krieger (1 December 2005). "Stanford Earns \$336 Million Off Google Stock". San Jose Mercury News, cited by redOrbit. Retrieved 2009-02-25.