Google PageRank-like algorithm dates back to 1941

February 19 2010, by Lisa Zyga

Since the 1940s, PageRank-like iterative algorithms have been used to rank industries, journals, and people.

(PhysOrg.com) -- When Sergey Brin and Larry Page developed their PageRank algorithm for ranking webpages in 1998, they certainly knew that the seeds of the algorithm had been sown long before that time, as is evident from their paper's references. But the Google founders may not have known just how far back PageRank's predecessors reach - nearly 70 years, according to Massimo Franceschet, who dug up a 1941 paper with a similar ranking method, as well as several other pre-Google papers with algorithms that show remarkable similarities to PageRank. Yet Brin and Page may have expected as much; after all, as Franceschet notes, the motto of Google Scholar is "Stand on the shoulders of giants."

In a recent study, Franceschet, a computer scientist at the University of Udine in Italy, has presented a brief history of iterative ranking methods

that predate PageRank. He also explains how the circular PageRank concept of determining the importance of a webpage based on the number of links it receives from important webpages, rather than by subjective expert evaluation, has provided an alternative way to define the quality of an item.

The 1941 predecessor of PageRank is a paper by the economist Wassily W. Leontief, who developed a method for ranking the values of a nation's various industrial sectors. Each industrial sector relies on the others, both for building materials (inputs) to manufacture its own products, and by selling its finished products (outputs) to other industries so they can manufacture their own products. Leontief developed an iterative method of valuing each industry based on the importance of the industries with which it is connected through input and outputs (similar to web links in PageRank). In 1973, Leontief earned the Nobel Prize in economics for his work in this area.

Other more recent PageRank-like algorithms have been used for ranking items in areas such as sociology and bibliometrics. In 1965, 33 years before Page and Brin developed PageRank, the sociologist Charles Hubbell published a method for ranking individuals. His premise was that "a person is important if it is endorsed by important people." Like PageRank and Leontief's algorithm, Hubbell's method is also iterative, with its outputs influencing its inputs, ad infinitum.

Later, in 1976, Gabriel Pinski and Francis Narin developed a journal ranking method in the field of bibliometrics. Here, the premise is that the importance of a journal is determined by the importance of the journals that cite it, which again uses the same circular reasoning as PageRank.

Most recently, the computer scientist Jon Kleinberg of Cornell University developed a ranking approach very similar to PageRank,

which was published around the same time of Brin and Page's publication (Brin and Page reference Kleinberg's work in their own paper). Kleinberg's method was also aimed at optimizing Web information retrieval. The algorithm, called Hypertext Induced Topic Search (HITS), referred to webpages as "hubs" and "authorities." These definitions are purely functional; hub pages point to authority pages, and authority pages are pointed to by hub pages. Mathematically, HITS is strikingly similar to PageRank, even though both were developed independently. Since they've been published, both papers have received widespread recognition and thousands of citations.

While PageRank has made [Google](#) a very powerful search engine, it had to radically reformulate the concept of quality to do so. The algorithm must constantly reevaluate each page as the importance of other pages varies - making quality seem fleeting, and no longer permanent.

"Expert evaluation, the judgment given by peer experts, is intrinsic, subjective, deep, slow and expensive," Franceschet writes. "By contrast, network evaluation, the assessment gauged [by] exploiting network topology, is extrinsic, democratic, superficial, fast and low-cost."

As Franceschet has shown, the new concept of value goes beyond webpages. Today, this "popularity contest" style of determining quality is stirring debate in academic circles in the area of research quality evaluation. Traditionally, evaluation of academic papers is done through expert peer review; the alternative is to use the PageRank-inspired Eigenfactor metric, which uses bibliometric indicators to evaluate research quality. Most likely, there will be other areas that see the use of PageRank-inspired methods redefining the concept of value.

  More information: Massimo Franceschet. ["PageRank: Stand on the shoulders of giants."](#) arxiv.org
via: [Technology Review](#)

© 2010 PhysOrg.com