The Potential for Democratizing Search Engines

DaNae M. Grubbs, Department of Computer Science
North Carolina Agricultural & Technical State University, Greensboro, NC
Professor Stephen Wicker, Department of Electrical & Computer Engineering
Cornell University, Ithaca, NY

The Internet is a global network that is home to mass amounts of information that is accessed daily. Search engines play a substantial role with this in regards to how the information is found. Search engines are also the quickest and most convenient way to find this information. Google, which is currently the number one search engine, ranks websites based on their relevancy and popularity using a math-based algorithm called "PageRank". The Internet is extremely massive in regards to the amount of websites that exist. Therefore, not every website is able to appear in the forefront of search engine results, which presents a dilemma. The websites that are at the forefront of the search engine results remain there for this very reason, because they are most "popular". It's not necessarily fair to have websites ranked according to a math algorithm that does not directly include user input. Fortunately, there have been attempts to engineer a search engine that would be more democratic. The Wikia Search engine was implemented to aid this problem. The search engine was community based and enabled users to rate, add and delete the search engine results. However, the search engine was only functional for 16 months. In this study, an analysis of the Wikia Search engine was performed to understand its features and analyze why it failed. Upon analysis, there was an overwhelming consensus that most search engine users are not interested in participating in the Wikia Search engine project and are ultimately content with the Google search engine.

INTRODUCTION

There is an extensive amount of websites that are present today. According to Pingdom.com, as of December 2009, 234 million websites are in existence. [1] Thousands of businesses and organizations have developed some form of a web presence, and with the introduction of personal blogs and websites millions of ordinary citizens have also followed the trend. With so many websites on the Internet, there is a great deal of competition due to the similarity in content of most sites. Websites are discovered through communication, linkage and search engines. Search engines are the most popular and convenient way for Internet users to discover web pages. Search engines are designed uniquely by their algorithm, which ultimately determines how search query results are ranked and displayed. According to Hitwise.com, the top three search engines today are Google, Yahoo! and Bing (formerly MSN Search). As of July 17, 2010, Google is ranked number one with a volume of 71.31%, dominating the search engines. [2]

The Internet is utilized as a resource for people to exercise their voice. However, not all of these "voices" are necessarily heard. Internet users have the ability and the right to create websites as they please, to produce the content they desire. One issue with the

Google Search Engine is that is doesn't necessarily grant the less popular sites the opportunity to appear on the initial page of the search engine results. How do these unpopular websites receive exposure from this? What would a more democratic search engine that allowed users to have an input on the search engine results look like?

The Wikia Search engine attempted to solve this problem. It allowed users to rate and annotate the websites that appeared in the search query results. Unfortunately, this search engine only lasted for approximately 16 months. Due to the Wikia Search engine's non-existence today, there was a major limitation of the research conducted. Without having the actual search engine to view and observe, a great deal of the information analyzed in this research is based on articles that feature facts and opinions of the search engine when it was available.

BACKGROUND

The goal of the world's largest search engine, Google, is to "organize the world's information and make it universally accessible and useful" and to create the "perfect search engine" that provides only intuitive, personalized, and relevant results.

Google founders Lawrence Page and Sergery Brin developed the algorithm as a part of their research project to develop a new kind of search engine. The actual Google "PageRank" algorithm is private, but details on how it ranks its pages have been released. People do have a general idea of how Google utilizes and has a trademark of the "PageRank" algorithm in which the patent is assigned to Stanford University.

Google defines "PageRank" as: "PageRank relies on the uniquely democratic nature of the web by using its vast link structure as an indicator of an individual page's value. In essence, Google interprets a link from page A to page B as a vote, by page A, for page B. But, Google looks at more than the sheer volume of votes, or links a page receives; it also analyzes the page that casts the vote. Votes cast by pages that are themselves "important" weigh more heavily and help to make other pages "important"." [3]

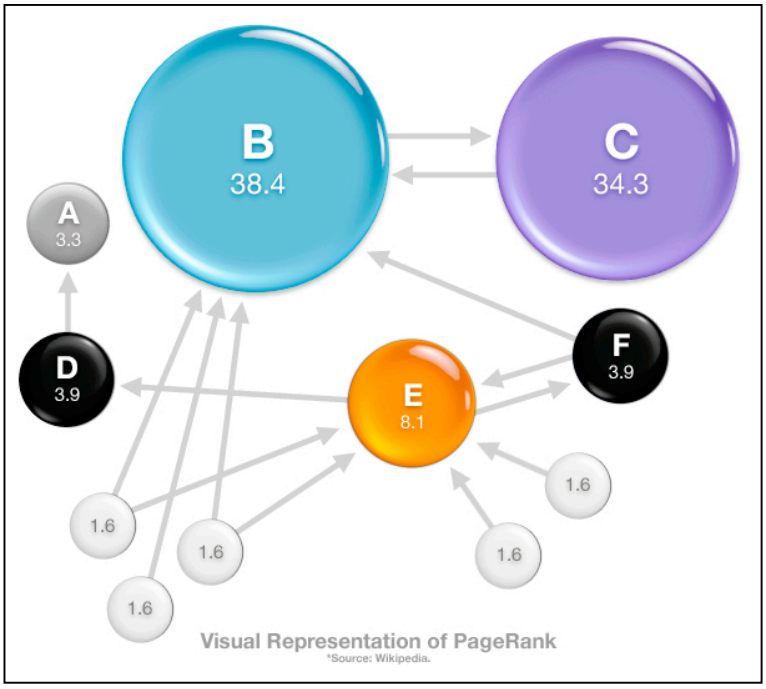

Figure 1: Web linkage structure of how the Google "PageRank" algorithm works. [4]

In Figure 1, "PageRank" is depicted by a web that shows inbound and outbound links. The circles represent a website and the number within the circle represents its particular weight. The weight is based on the number of inbound links to the website. The websites that are weighted the least provides outbound links to

the websites that are weighted the most. On the contrary, the websites that are weighted the most have a high number of inbound links, and therefore appear at the forefront of search engine results.

Jimmy Wales, the founder of Wikipedia, implemented the Wikia Search engine to give users the ability to control the search engine results. Wikipedia is an online community-based Encyclopedia that allows users within the community to contribute to it. The site has spanned tremendously over the past four years. With the success of Wikipedia, Wales had a similar concept in mind for the implementation of the Wikia Search engine.

The Wikia Search engine allowed users to edit the results that appeared in the results. Once revised, the edits were instantly visible for everyone to view. Users had the ability to rate the results from 1-5 stars, with 5 being the highest. Users were also able to delete, add and suggest related results. Users who participated in the Wikia Search engine had the option of registering on the search engine's website. This ultimately means that anyone could contribute, and information about their identity would remain anonymous. The Wikia Search engine also had a social networking component. It allowed members to register and create a profile. With a profile, members were able to upload photos, send private messages to other users, post comments on their profile and be notified of other users' actions.

A little over a year after the search engine launched, Wales shut down the search engine. Wikia Search had reportedly only inherited 10,000 users per month.[5] Wales still holds out hope that community-based searching will once again become an online reality. As he stated in his blog, "I will return to again and again in my career to search, either as an investor, a contributor, a donor, or a cheerleader."

ANALYSIS

A survey conducted by CNET.com shows the results of what people initially thought the Wikia Search Engine would look like (see fig. 2). 64.7% of the survey participants thought the search engine would be chaotic and depict unstable search engine results. If people had this preconceived notion about the search engine, then they were more than likely not interested in using it. This analysis supports the statistic that the site only had a unique number of about 10,000 users per month.[5] The 6.9% of the survey participants who thought the search engine would not make a difference also are likely to be the users who did not contribute to this project. In contrast, it is the 23.2% of the users who thought the search engine would produce better results are more than likely the users who contributed to the Wikia Search engine.

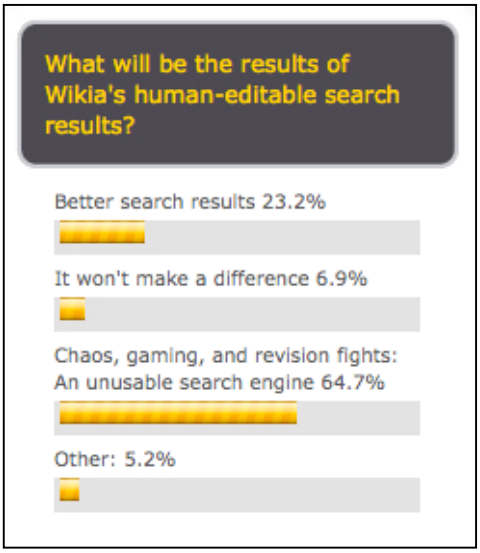

Figure 2: CNET.com survey outcome of what people initially thought the Wikia Search Engine results would look like. [6]

The concept of having a transparent search engine is keen, yet risky. The issue in this is that a transparent search engine equates to a very vulnerable search engine. A search engine of such nature is highly susceptible to frequent spamming of users anticipating to have their particular website appear at the forefront of the search results displayed.

Registration on the Wikia Search engine website was indeed optional, which ultimately attracted users. However, if people have the option of signing up for a service, they are not likely to do so because of privacy concerns and time commitment. Not everyone is willing to reveal their identity or devote the time to register. With the Wikia Search engine having social networking features, Wales believed he could recruit and retain users, which would thereon lead users to frequent contribution.

CONCLUSION

The Wikia Search Engine had potential to be great but many search engine users were just simply not interested. When it comes to search engines, Google is simply the prevalent choice. Many people who utilize the Internet are content with the Google search engine because it simply does what they want it to do—produce relevant results.

Google is not perfect but it is very efficient at producing its reliable results. If Google were to modify its search engine to sample websites that are in the medium of the results, some of its hierarchy concerns would be addressed. It would allow users to see results that are not only at the forefront, but also those that lie in the median of the results. The Google search engine could achieve this by showing the top three results, and then the remaining results displayed would be two results that are typically shown on page 10, 100, 1000, etc. If Google could display these results on the first page, it would allow users to get a more broadened search experience while allowing “underdog” websites to get more exposure.

Future work of this research would involve looking out for the re-release of the Wikia Search engine. Wales believes that the search engine will one day dominate the industry and be Google’s superior. When the search engine debuts again, an observation of its structure will be carried out to determine if it has changed any of its previous features and the likelihood of more users' participation.

REFERENCES

[1] "Internet 2009 in numbers." *Royal Pingdom*. N.p., 22 Jan 2010. Web. 24 Jul 2010. <http://royal.pingdom.com/2010/01/22/internet-2009-in-numbers/>.

[2] "The Power of Competitive Intelligence." Hitwise United States. N.p., 17 Jul 2010. Web. 29 Jul 2010. <http://www.hitwise.com/us/datacenter/main/dashboard-10133.html>.

[3] Bury, Tony. "SitePoint." *Google PageRank - Democracy or Corporate Muscle?*. N.p., 16 Jan 2003. Web. 17 Jul 2010. <http://articles.sitepoint.com/article/democracy-corporate-muscle>.

[4] Pendery, Rafferty. "Simple Explanation of How PageRank Works." Rafferty Pendery. Rafferty Pendery, n.d. Web. 24 Jul 2010. <http://www.raffertypendery.com/internet-marketing-articles/page-rank.php>.

[5] Needleman, Rafe. "Wales giving up on Wikia Search." Webware- CNET. N.p., 31 Mar 2009. Web. 25 Jul 2010. <http://news.cnet.com/8301-17939_109-10207896-2.html>.

[6] Needleman, Rafe. "Wikia Search launches the hackable search engine." Webware-CNET. N.p., 03 Jun 2008. Web. 24 Jul 2010. <http://news.cnet.com/8301-17939_109-9958036-2.html>.

ACKNOWLEDGMENTS

This research is support by the Team for Research in Ubiquitous Secure Technology (TRUST) Research Experience for Undergraduates. Thanks to Dr. Stephen Wicker of the School Electrical & Computer Engineering and graduate student, Nathan Karst.