

Analysis of Search Engine Optimization Algorithm

Mehzabul Hoque Nahid

Lecturer, MIS, American International University-Bangladesh, Dhaka, Bangladesh

Abstract: SEO stands for “search engine optimization.” It is the process of getting traffic from the free, organic, editorial or natural search results on search engines. The earlier (or higher ranked on the search results page), and more frequently a site appears in the search results list, the more visitors it will receive from the search engine's users, and these visitors can be converted into customers. SEO may target different kinds of search, including image search, local search, video search, academic search, news search and industry-specific vertical search engines. As an Internet marketing strategy, SEO considers how search engines work, what people search for, the actual search terms or keywords typed into search engines and which search engines are preferred by their targeted audience. Optimizing a website may involve editing its content, HTML and associated coding to both increase its relevance to specific keywords and to remove barriers to the indexing activities of search engines. Promoting a site to increase the number of backlinks, or inbound links, is another SEO tactic. As of May 2015, mobile search has finally surpassed desktop search, Google is developing and pushing mobile search as the future in all of its products and many brands are beginning to take a different approach on their internet strategies.

Keywords: SEO, HTML, Factors, Volume, Crawl.

I. INTRODUCTION

SEO is a marketing discipline focused on growing visibility in organic (non-paid) search engine results. SEO encompasses both the technical and creative elements required to improve rankings, drive traffic, and increase awareness in search engines. There are many aspects to SEO, from the words on your page to the way other sites link to you on the web. Sometimes SEO is simply a matter of making sure your site is structured in a way that search engines understand. SEO isn't just about building search engine-friendly websites. It's about making your site better for people too. The main objective for creating a website by any company is to make aware of about the company and its products / services [1] [3]. For this, Internet marketing plays an important role in driving the visitors to site using SEO techniques. Now, it becomes essential that the site should be listed in the top pages of the search engines and upbeat the competitor's site.

The process of getting search engines to pay attention to your web site requires specific, personalized service. To achieve this, proper SEO tools and techniques are applied to the site to achieve higher rankings and make site search engine friendly. Although the basic principle of operation of all search engines is the same, the minor differences between them lead to major changes in results relevancy. For different search engines different factors are important. There were times, when SEO experts joked that the algorithms of Bing are intentionally made just the opposite of those of Google. While this might have a grain of truth, it is a matter of fact that the major search engines like different stuff and if you plan to conquer more than one of them, you need to optimize carefully. There are many examples of the differences between search engines. For instance, for Yahoo! and Bing, on-page keyword

factors are of primary importance, while for Google links are very, very important. Also, for Google sites are like wine – the older, the better, while Yahoo! generally has no expressed preference towards sites and domains with tradition (i.e. older ones). Thus you might need more time till your site gets mature to be admitted to the top in Google, than in Yahoo!

II. HOW SEARCH ENGINE WORKS

The first basic truth you need to know to learn SEO is that search engines are not humans. While this might be obvious for everybody, the differences between how humans and search engines view web pages aren't. Unlike humans, search engines are text-driven. Although technology advances rapidly, search engines are far from intelligent creatures that can feel the beauty of a cool design or enjoy the sounds and movement in movies. Instead, search engines crawl the Web [2], looking at particular site items (mainly text) to get an idea what a site is about. This brief explanation is not the most precise because as we will see next, search engines perform several activities in order to deliver search results – crawling, indexing, processing, calculating relevancy, and retrieving.

First, search engines crawl the Web to see what is there. This task is performed by a piece of software, called a crawler or a spider (or Google bot, as is the case with Google). Spiders follow links from one page to another and index everything they find on their way. Having in mind the number of pages on the Web (over 20 billion), it is impossible for a spider to visit a site daily just to see if a new page has appeared or if an existing page has been

modified, sometimes crawlers may not end up visiting your site for a month or two. What we can do is to check what a crawler sees from our site. As already mentioned, crawlers are not humans and they do not see images, Flash movies, JavaScript, frames, password-protected pages and directories, so if you have tons of these on your site, you'd better run the Spider Simulator below to see if these goodies are viewable by the spider. If they are not viewable, they will not be spidered, not indexed, not processed, etc. - in a word they will be non-existent for search engines. After a page is crawled, the next step is to index its content. The indexed page is stored in a giant database, from where it can later be retrieved. Essentially, the process of indexing is identifying the words and expressions that best describe the page and assigning the page to particular keywords. For a human it will not be possible to process such amounts of information but generally search engines deal just fine with this task. Sometimes they might not get the meaning of a page right but if you help them by optimizing it, it will be easier for them to classify your pages correctly and for you – to get higher rankings. When a search request comes, the search engine processes it – i.e. it compares the search string in the search request with the indexed pages in the database. Since it is likely that more than one page (practically it is millions of pages) contains the search string, the search engine starts calculating the relevancy of each of the pages in its index with the search string. There are various algorithms to calculate relevancy. Each of these algorithms has different relative weights for common factors like keyword density, links, or met tags. That is why different search engines give different search results pages for the same search string. What is more, it is a known fact that all major search engines, like Yahoo!, Google, Bing, etc. periodically change their algorithms and if you want to keep at the top, you also need to adapt your pages to the latest changes. This is one reason (the other is your competitors) to devote permanent efforts to SEO, if you'd like to be at the top.

The last step in search engines' activity is retrieving the results. Basically, it is nothing more than simply displaying them in the browser – i.e. the endless pages of search results that are sorted from the most relevant to the least relevant sites.

III. SEO CONTENTS FOR SEARCH ENGINES

SEO no longer means scattering keywords like Hansel and Gretel throwing breadcrumbs. The newest search engines scan pages almost as your readers might. Jakob Nielsen, a researcher and expert in human-machine interaction at the Technical University of Copenhagen, found that almost 80 percent of a web site's visitors scanned the page rather than reading it line by line. They spent their first fractions of a second on the page deciding if it was worth their time. Search engine programmers still use this research to devise algorithms that provide more organic and meaningful rankings. The same things that catch a visitor's eye will get a search engine's attention. The upper left

corner of the page is the most valuable real estate on the page, as it's where a reader's eyes go first. Put important text there so search engines and people will see it immediately. It's also a good spot for boxed text and itemized lists, both of which appeal equally to carbon-based and silicon-based brains. Bold text makes people and machines notice, but use those tags judiciously. Too much bold text looks like an advertisement and will cause search engines to devalue your site. Italic text bold HTML tags should surround meaningful concepts, not emphasis words. Bolding a "very" or italicizing a "more" means nothing to a search engine, so apply those tags to important concepts and sub-headings. Searches now look for associated terms and relevant phrases, not just keywords. A person picks up meaning from context and readily distinguishes the term "clipping" as it applies to hair from the same word as it refers to film stock or video game graphics. Let your visitors -- human and machine -- know whether you're talking about German shepherds as a dog breed or as an exciting career in European wool and mutton [4]. In your SEO text, include synonyms and relevant terms to let search engines recognize the purpose of your site. Happily, there's a way to work these terms into our content without monitoring keyword and key phrase percentages: simply write the kind of engaging copy that people like to read. If we write for readers, the search engines will follow.

IV. SEO SUCCESS FACTORS

There are three major groups covered by Search Engine Land's Periodic Table of SEO Success Factors:

- On-The-Page SEO
- Off-The-Page SEO
- Violations

Within each group are subgroups, as each chapter of this SEO guide will explain. These subgroups contain one or more individual SEO factors with a specific weight or importance. Violations, while a group unto themselves, are displayed under the group and subgroup to which they're associated.



Fig1: SEO factors Periodic Table

Those two letter acronyms you see on the chart above? That's our play on the periodic table of elements and the

letter representations, or symbol [5], of each element. You may have had to remember that the symbol for gold was Au or that iron's was Fe. In the Periodic Table of SEO, we've tried to make it slightly more intuitive. The first letter of each "SEO element" comes from the subgroup that it's in and the second letter stands for the individual factor.

V. SEO PROCEDURE

- A. Initial Setup Strategies
- B. Keyword and Competitive Analysis
- C. Site Optimization
 - 1. On Page Site Optimization
 - A. Keyword identification per page
 - B. Content improvement as per keyword targeted
 - C. Title / Meta Tags and description
 - D. File naming
 - E. Web Page Analysis
 - 2. Off Page Site Optimization
 - A. Link Building
 - D. Paid Marketing / Promotion
 - 1. Pay Per Click Campaign
 - 2. Affiliate Marketing
 - E. Site Verification & Positioning Reports



Fig 2: SEO Lifecycle

VI. SEO SEARCH MARKETING LANDSCAPE

No single SEO factor will guarantee search engine rankings. Having a great HTML title won't help if a page has low quality content. Having many links won't help if they are all low in quality. Having several positive factors can increase the odds of success while the presence of negative factors can worsen those odds [6].

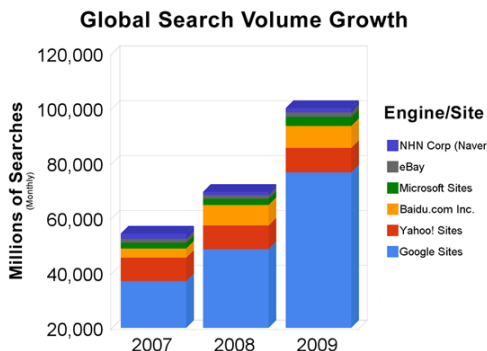


Fig 3: SEO Volume Growth up to 2009

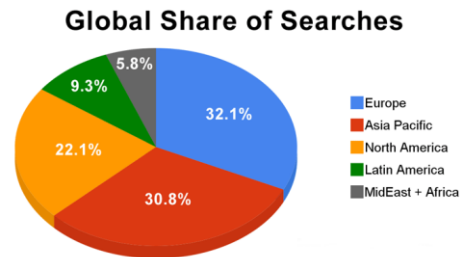


Fig 4: Searches History

To avoid undesirable content in the search indexes, webmasters can instruct spiders not to crawl certain files or directories through the standard robots.txt file in the root directory of the domain. Additionally, a page can be explicitly excluded from a search engine's database by using a meta tag specific to robots [7]. When a search engine visits a site, the robots.txt located in the root directory is the first file crawled. The robots.txt file is then parsed, and will instruct the robot as to which pages are not to be crawled. As a search engine crawler may keep a cached copy of this file, it may on occasion crawl pages a webmaster does not wish crawled [8]. Pages typically prevented from being crawled include login specific pages such as shopping carts and user-specific content such as search results from internal searches. In March 2007, Google warned webmasters that they should prevent indexing of internal search results because those pages are considered search spam.

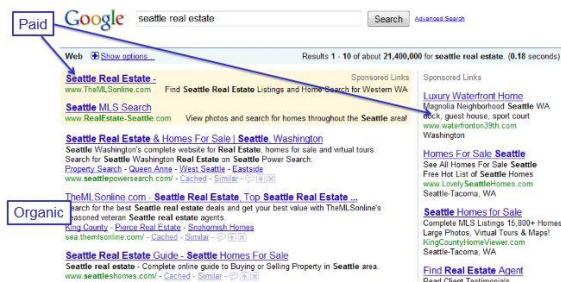


Fig 5: Organic Vs Paid Search

Result Position	Average Percentage of Clicks	Fraction of Top Ranking Result
#1	42.1%	N/A
#2	11.9%	3.5X Less
#3	8.5%	4.9X Less
#4	6.1%	6.9X Less
#5	4.9%	8.5X Less
#6	4.1%	10.4X Less
#7	3.4%	12.3X Less
#8	3.0%	14.0X Less
#9	2.8%	14.8X Less
#10	3.0%	14.1X Less
#11-20	10.1%	N/A
#21-1,000	1.2%	N/A

Fig 6: Click through Rates

VII. CONCLUSION

The availability of information on the internet is growing dramatically. But it will be useful if the user finds the correct information in less amount of time. So in order to stand out from the competition means in order to increase the sites visibility, Search Engine Optimization (SEO) plays an important role.

SEO is cost-effective process to promote internet marketing. It helps to bring the customers to the website. SEO can be done by two ways- on page and off page SEO. On-Page optimization can be done on the page like to check broken tags, links, checking Meta tags, Alt tags etc. Off-Page optimization means giving quality back links to the particular web page to improve the website ranking.

SEO is a process consist of various stages keyword research, goal setting, page optimization, content development, link building and then to follow-up the entire process.

This paper has shown SEO research that has done to increase website ranking and thus to fetch the website traffic. So this paper can be used as basis for SEO engineers. In the upcoming years, a simple approach is not enough.

Efficient SEO requires managing three basic pillars- the content, the keyword and the link building. With the help of these three pillars one can create basic SEO structure that will support high ranking of the web pages in the coming future.

ACKNOWLEDGMENT

I would like to show my gratitude to **Ahmed Abdullah**, Lecturer of RUD for sharing his knowledge and expertise on SEO with me during the course of writing this article. I also would like to thank **Dr.Manzur Ashraf** for his insight that greatly assisted in my research work.

REFERENCES

- [1]. S.Mukherjee, "A probabilistic model for optimal searching of the deep Web", 2003.
- [2]. Fuxue Wang; Yi Li; Yiwen Zhang; Coll. of Econ. & manage, "An empirical study on the search engine optimization technique and its outcomes," Artificial Intelligence, Management Science and Electronic Commerce AIMSEC), Aug. 2011.
- [3]. Ma Xiao ling, Wu Yong he." For research of search engine optimization (SEO)". Information Retrieval. 2005, 12(1):6~ 10
- [4]. Graf, Cousin. Search Engine Optimization. Beijing: Qinghua University Press, 2007.
- [5]. Liu Xiaolin." Chinese search engine key technologies". Hubei University of Technology Reported.2006, 6(1):9~12
- [6]. Yi Dan. "Search engine optimization (SEO) Research and Implementation". Petroleum Industry Computer Applications. 2006, (1):132~ 134
- [7]. Grappone J, Couzin G. Search Engine Optimization: An Hour a Day. Yang Mingjun translation. Beijing: Tsrnghua University Press, 2 007.
- [8]. Fei Wei, Huang Ruhua. "Based on user behavior analysis of search engine optimization strategy". Library and Information Service, 2005 (10):314~320

BIOGRAPHY



Mehzabul Hoque Nahid is a lecturer of Management Information System department at American International University-Bangladesh. He has obtained his Masters of Information Technology degree from Swinburne University of Technology Australia in 2011. He completed his bachelor degree from American international university Bangladesh (AIUB) in 2007. His research interests include E-governance, SEO, data mining, professional issues in IT, human computer interaction, knowledge management, ICT in Agriculture, network management, network modeling and simulation.