

Past, Present and Future of Search Engine Optimization

Saeid Saberi¹, Golnaz Saberi² and Masnizah Mohd¹

¹Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, 43600 Selangor, Malaysia

²Faculty of Applied Information Technology, Chalmers University of Technology, 412 96 Göteborg, Sweden

s.saberi@yahoo.com newg_s@yahoo.com mas@ftsm.ukm.my

ABSTRACT

Today the role of Internet in all aspects of our lives has become more and more vital. Accordingly, for businesses to grow, being active and visible on the Internet has become a new requirement. On the other hand, based on confirmed study results, the most important gateways to web are search engines. The significance of search engines has led to foundation of a new field of research: Search Engine Optimization (SEO). In the present study we have examined the previous, current and prospective future approaches in SEO and concluded that it is an important requirement for success of websites to be considered. In addition, we have offered a number of suggestions for webmasters who are interested in the latest SEO strategies.

KEYWORDS

search engine, search engine optimization, seo, website visibility, website optimization, keyword optimization, structure optimization, content optimization, quality of content, social media optimization.

1 INTRODUCTION

Today, search engines occupy a prominent position in the online world. According to StatOWL 2012, approximately 82% of users enter the Internet environment via search engines (www.statowl.com/network_visitor_source.php). Therefore, to get high traffic, websites need to be ranked highly by search engines and placed on the first page or first few rows of search results. Due to this fact webmasters either pay the search engines to get internet advertisement or choose a more natural approach: Search Engine Optimization (SEO), developing websites or web

pages in a way that they are found easily and ranked highly by search engines. Due to lower costs, “SEO” is the preferred choice over Internet advertisement.

SEO is totally based on the algorithms that are used for searching and indexing websites by search engines [1]. Therefore, examination of its methods naturally involves a lot of research about search algorithms. This study divides the development of SEO to three parts: Past, Present and Future. Past refers to the period during which SEO methods were mainly based on the shortcomings of search algorithms; these methods were mainly unethical. But later as algorithms improved and became intelligent, search engines became able to catch the websites using them which would get penalized or totally removed from the index. The second part of this research, Present, is devoted to the study of the most important state-of-the-art techniques that are employed by SEO specialists. The correct application of these methods can help websites to get high ranks for certain keywords. However, search algorithms are continuously being revised towards being more and more intelligent. In order to hold a position on top of the search lists, webmasters must continually monitor the on-going improvements and adjustments. In the Future part we study the possible orientation of enhancement of search algorithms and correlate them with future of SEO.

2 SEARCH ENGINE

Typically, search engines are software programs employed for data collection about websites on the Internet. The main tasks of a search engine

include: page collection, page analysis and page sorting. In the following part of this chapter we review these tasks briefly.

Page collection refers to the activity of collection and storing data from Internet into high-capacity databases of search engines. Such a database is referred to as an *index* and the operation is called *indexing* [2]. A new web page is added to index by spider software of a search engine which utilizes the hyperlink structure of known web pages to discover new ones. Spiders work all the time and update the indexed web pages frequently. An index typically includes several types of data about websites/web pages: URL, keywords, HTML code, and related links.

During *page analysis* the collected web pages undergo a series of analysis and processing, mainly including label filtration, extraction of web page text information, text parsing, and establishment of an index between keywords and the web page.

Finally, when a user enters a search query, it is compared to the data in the index and a sorted list of relevant results regarding the query is rendered by the search engine [2]. The process of ranking the results based on their relevance to the query is called *page sorting*, which is a critical competitive factor in search algorithms.

3 SEARCH ENGINE OPTIMIZATION

A study in 2013 shows that 99.57% of Internet users have only tried the first page of search engine results (www.statowl.com/search_engine_results_page.php). Therefore, the position of an item on the result list, which is determined by its rank, is a highly influential factor on the related web page traffic. Usually, webmasters use particular techniques so that when web pages are indexed by search engines, they get high ranks for certain keywords; this process is called Search Engine Optimization (SEO). SEO techniques may be considered in website coding, presentation, structure, marketing strategies, etc.

4 THE PAST OF SEO

In the early days of search engines, many webmasters discovered ways to get high ranks by artificial and unethical manipulation of search engines, the so called ‘black hat’ techniques. At that time search algorithms were quite transparent and it was not so difficult to ‘fool’ the search engines to assign high ranks to pages which did not really ‘deserve’ it. But later search engines started to fight back by using specific algorithms, filters and penalties to stop spam websites from being indexed. Today search engines penalize websites which use such methods by assigning them low ranks or completely removing them from the index. However, still many websites use black hat techniques and occasionally succeed to get high ranks by accident. Some of these techniques are as follows [1], [3]:

- *Keyword stuffing*: The most common trick during which the page is loaded with numerous instances of the given keyword to fool the search engine to consider them as relevant content.
- *Link farm*: A series of websites created solely for housing gratuitous links to a collection of websites or web pages. It can also be a network of websites interlinking each other. The purpose of a link farm is to increase the number of incoming links of other web pages.
- *Hidden Text*: Text with the same colour as the background or very close to it which website’s visitors cannot see, but is still readable by the search engines.
- *Cloaking*: Presenting content to search engines that is different from that which human visitors see.
- *Doorway Pages*: Pages with high ranks for certain keywords or phrases but no real valuable information. Usually many of these pages have link to one particular web page with unrelated content to visitors’ queries.
- *Redirects*: A very common tactic that has been brought in as a compliment to doorway pages. Such pages automatically redirect to the target web page.

- *Interlinking*: Building multiple websites and linking them together to raise the overall link popularity of each of them. This trick is the most difficult to detect when applied "correctly".

5 THE PRESENT OF SEO

SEO techniques are observed from four aspects: keywords optimization, structure optimization, content optimization and link optimization [4]. This section encompasses a relatively detailed explanation of these aspects.

5.1 Keyword Optimization

Users mainly apply keywords as search query to find web pages, while search engines use them as search criterion to retrieve web pages from their index. Therefore, keyword optimization is the primary part of SEO. During keyword optimization, the core keywords of a website are determined at first, secondly keyword phrases and terms are generated around those core keywords and finally they are appropriately distributed in the web page's text.

5.1.1 Keyword Selection

Choosing the proper keywords is one of the most demanding steps in SEO. The main principles of the keyword selection are [5]:

- *Related Content*: Keywords must be associated with the website content. Even though a big number of keywords can bring about a lot of search hits, but if they are not well content related the traffic would not be converted into the actual utilization of the webpage. This traffic would actually become a burden on the website because it consumes bandwidth.
- *Large Quantity of Searches with Less Competition*: The best keywords are those which are most frequently searched for but least competed against. However, the majority of common keywords are very competitive. Although it is not very easy, but keywords with relatively more searches and less competition can still be found through a sufficiently detailed

and extensive research. A useful tool for discovering potential keywords is Google's Keyword Tool, which returns not only the estimated search volume of a keyword but also hundreds of related keywords and phrases, in particular marketable ones.

- *Keywords should not be too general*: If keywords are too general, competitive, and unclear in the search intent, the conversion rate will be reduced. For example, a clothing company might assume the word 'clothing' as a good keyword. However, this word as well as most of the general industry terms is overly broad. If such words are used as main keywords, they would either have no positive effect on the website's traffic or even if the traffic is increased, the conversion rate still remains very low.
- *Keywords should not be too specific*: a very specific word is rarely applied as a search query by users.

5.1.2 Keyword Density

Keyword density refers to the ratio of total number of times that a keyword appears on a web page to the total number of words on the page. In general, from the point of view of a search engine a higher keyword density on a web page indicates more relevance between the web page and the keyword. However, this factor must be applied with care, as keyword stuffing is classified as malicious behaviour, and there is a risk of being punished. A range of 3%-8% is considered as the most efficient keyword density by SEO specialist [4].

5.1.3 Keyword Distribution

The location of keywords in a web page is an important issue in SEO. Webpage title, Meta keywords section, Meta Description section, page header tags (H1,-H6), anchor texts of internal and external links and Alt attribute part of graphs, are the most suitable positions that can contain keywords. In addition, as the text content is read sequentially, from top to bottom and from left to right, there can be 'focus' points in the text where the keywords are placed; such as the top part and

left side of the page, title, and the section within the first 200 words of the main text.

5.1.4 Domain name optimization

Domain name is considered as the most reliable factor by search engines. Due to this fact, the selection of domain name is a determinant factor in SEO. For an English website, domain name and web page file names can involve keywords. For instance, embedding the keyword 'video' in the domain name would be an inherent advantage for the whole website considering queries involving this keyword. Moreover, search engines give higher priorities to domain name suffixes that are used for a specific purpose. For example '.edu', '.gov', '.org' and '.net' are prioritized over the general-purpose suffix '.com'.

5.2 Structure Optimization

Structural optimization of a website ensures that viewers and search engines are able to find website's content more quickly and easily. The website structure must be clear and the depth of the general directory must be limited to a maximum of four layers. Moreover, websites should avoid imperceptible navigation system. For example, some web pages use frames and JavaScript or Flash buttons which look stylish but search engines cannot detect them. Generally speaking, search engines ignore image, Flash, and JavaScript code, so the major focus must be on the HTML code and text content of web pages.

Another technique for optimizing website structure is creating and submitting a sitemap which includes URLs that may not be discoverable by search engines' normal crawling process. It helps to make sure that the whole website is visible for search engines.

The third important factor in structure of a website is Page Size: The sum of file sizes of all the elements that make up a web page, including the defining HTML file as well as all embedded objects. Most search engines do not index pages that are greater than a certain size fully, so it is important to set a limit for the page size.

And last but not least is the Title tag is one of the most important components in search results. Most search engines assume that the words in the title tags have a close correlation with their proceeding text. Furthermore, apart from being informative a good title must be brief as well so the users can guess the right subject at the first glance. That is why it is recommended to keep the title length below 80 characters.

5.3 Content Optimization

Search engines love unique and original content. If the content of a website is only web pages copied from other websites, most likely it will not be indexed by search engines. In addition, constant update is the most basic condition for survival and development of a website. The frequency with which the index entry of a website is updated is determined by the update frequency of the website. If the website content is frequently updated, then search engine spiders will visit it more often and grab more pages.

5.4 Link Optimization

5.4.1 Incoming Links

For a website to get a higher rank it is essential to have as much high quality incoming links as possible. Therefore, it is worthwhile to use a link analysis tool to analyse the linkage status of the website. When a spider scans the source HTML of a webpage it follows the links to external web pages and indexes them as well. Thus, it will be more beneficial for a new website to receive high quality links rather than submitting itself to the search engine.

5.4.2 Outgoing Links

An Outgoing link is a link to another website. These links are also analysed by search engines. A link to an external website with a close subject to the original one is considered as a positive point [6]. While if the external website is spam or of low quality the rank of the original website would be affected negatively. Therefore, the outgoing links should be contained with care.

5.4.3 Internal Links

The web pages within a website should also be properly linked to each other. Apart from facilitating the navigation through website's pages and increasing the overall traffic, a complete internal link structure is also an asset for improving the rank of internal web pages [6]. From the perspective of SEO, it is important that each page has a link to the home page and the link structure is examined frequently to prevent broken links.

6 THE FUTURE OF SEO

Today, technology goes much faster than humans, in other words the 'future' is happening much sooner than what people expect. Therefore, it is very hard to talk about the future of some technology, especially in the world of computers, in the long run. However, we can extrapolate from the present course of development of a technology, like search engines, to find out the most probable direction of its future path. In the case of SEO we need to study the shortcomings and strength of search algorithms to be able to recognize possible enhancements or changes that might be made in the future. Here we present 4 issues that we believe are the most influential factors that are and will be determinant in how successfully a website can increase its traffic through SEO.

6.1 Quality of Content

This might sound intuitive that good quality is the strategy that works best in the long run. But how is it related to SEO? The answer is that a search engine is also a product developed for raising money. To be marketable a search engine must offer good quality, which is mainly based on the quality of its search results. The more relevant and up-to-date are the offered results the more satisfied are the users of search engines. Therefore, it is natural that search engine developers dedicate a great deal of their effort to prune out unrelated or low quality web pages from the highly ranked results. Accordingly, the first

winning factor in SEO is and will be high quality content.

6.2 Natural back-links

In Chapter 4 the influence of incoming links on the rank of websites was discussed. As mentioned before, receiving incoming links from popular websites can be even more effective than submitting the website to the search engines. The best strategy to get good reputation is publishing original and excellent content. High quality content sells best not only for its natural effect on the website's rank but also because of its capability to attract links from popular websites. Moreover, it can motivate readers to share the content on social websites and thus making the original website more well-known which could contribute to attracting back-links.

6.3 Popularity

The third factor that we believe will become more important in the near future is user's interests. Today shopping websites usually have a recommendation system that offers relevant products to the viewers besides the one that they are looking for. One way of selecting products to offer is identifying users with similar interests and recommending them the products that the others have purchased. Now some search engines save cookies on user's computer to 'memorize' their previous choices and sort the next search results according to this 'history'. So, why not use this method as a profile system for identifying similar users? We believe that it is highly probable that a similar strategy be used by search engines in the ranking process in the future. In that case, the more visitors attracts a website the more it is recommended to other users with similar profiles. That is another reason that webmasters must try to expose the website to as many users as possible.

One way to increase the website's popularity, that is already becoming very common, is having profile on social networking websites, which can bring about a broad network of users. Furthermore, due to the fact that internet is increasingly being used on alternative platforms

other than PCs or laptops, like smart phones, tablets, televisions, etc, it is crucial for the websites to have other versions that are adaptable with these popular platforms as well. In addition webmasters must keep an eye on the alternative technologies, like Apple's 'Siri', that will probably replace the old in the internet. For instance it is a good idea to build a profile on Google Places or Yelp that are likely to be utilized by such third-party programs.

6.4 Activity on Social Media

As mentioned before, being active on social media can indirectly build up website's rank by providing it with more popularity. However, activity on social websites can be directly efficient as well for search engines have recently shown a tendency to use social websites as another source of information [7]. In 2010, it was confirmed that Google and Bing consider links and reputation and authority of linking author on social media, specifically Twitter and Facebook, as another factor in their ranking process. On the other hand, as social activity on the web shows no signal of ceasing, it is highly probable that these 'social signals' receive more attention from search engines in the future.

In relation to the above discussion, it is worthwhile to mention the concept of Social Media Optimization (SMO). As stated by marketer Rohit Bhargava in 2006, SMO is the practice of "optimizing a site so that it is more easily linked to, more highly visible in social media searches on custom search engines, and more frequently included in relevant posts on blogs, podcasts and vlogs". As it is evident from the above definition, SMO techniques help to increase a website's visibility in social media as well as provide it with more incoming links from relevant sources which are both assets to website's rank by search engines. Consequently, SMO techniques can be considered as part of SEO.

7 CONCLUSION

Search Engine Optimization (SEO) strategies are closely coupled with the search algorithms applied

by search engines. In the past, SEO techniques were mainly based on the flaws of search algorithms which caused many spam websites to get the chance to increase their rank by using 'black hat' techniques. Nevertheless, in time search engines optimized their algorithms to be able to catch and penalize fake content. Right now, although not being capable of identifying each and every instance of spamming, search engines are able to weed out most of such cases. Furthermore, search engines consider several factors as being of great significance when ranking websites. A number of these factors include keyword selection, website structure, quality and originality of content and incoming, outgoing and internal links which must be taken into account by webmasters. Finally, we conclude by suggesting several factors as being determinant in the future of SEO: quality of content, natural back-links, popularity among visitors and maintaining an active profile on social media.

8 REFERENCES

1. Killoran, J.B.: How to Use Search Engine Optimization Techniques to Increase Website Visibility. In: Professional Communication, IEEE Transactions on , vol.56, no.1, pp. 50-67 (2013)
2. Yalçın, N., Köse, U.: What is search engine optimization: SEO?. In: Procedia - Social and Behavioral Sciences – Proceeding of the World Conference on Learning, Teaching and Administration Papers, pp. 487-493 (2010).
3. Ao-Jan, S., Hu, Y., Kuzmanovic, A., Cheng-Kok K.: How to Improve Your Google Ranking: Myths and Reality. In: IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT), pp. 50-57, IL, USA (2010).
4. Hui, Z., Shigang, Q., Jinhua L., Jianli, C.: Study on Website Search Engine Optimization. In: International Conference on Computer Science and Service System, pp. 930-933, Nanjing, China (2012).
5. Wang, F., Li, Y., Zhang, Y. An empirical study on the search engine optimization technique and its outcomes. In: 2nd International Conference on Management Science and Electronic Commerce (AIMSEC), pp. 2767-2770, Chongqing, China (2011).
6. Zhang, L., Zhang, J., Ju, Y.: The research on Search Engine Optimization based on Six Sigma Management. In: 2011 International Conference on E-Business and E-Government (ICEE), pp.1-4, Beijing, China (2011).
7. Bhargava, R.: 5 rules of social media optimization (SMO). Influential marketing blog. [Online]. Available:

<http://www.rohitbargava.com/2010/08/the-5-new-rules-of-social-media-optimization-smo.html>