

SURVEY ON IMPORTANCE OF SEARCH ENGINE OPTIMIZATION & It's TECHNIQUES

K M Patel¹, Prof. Rajesh Pathak²

^{1,2}*Computer Science Engineering, Monad University, (India)*

ABSTRACT

In this research survey, author has covers previous research done on the topic of SEO, search engine fundamentals, and a brief overview of the history of the Internet. This will help give a much better understanding of the importance of SEO in the current state of the Internet and in information search. Moreover, an analysis of the importance of becoming relevant and maintaining visibility among a large number of websites will also be discussed. This survey divides into various sections as follows: Section-I will cover related research in the SEO field. Section-II describes important events that took place for the development and history of the Internet, early browsers, the dramatic growth of the Internet and rise of search engines. Section-III provides a brief summary of the beginnings of Google, how it ranks pages using its proprietary PageRank algorithm and the importance of hyperlinks in the scheme of the Internet. Section-IV discusses the history of SEO and its goals, the different types of SEO implementations (on-page SEO and off-page SEO), and the difference between white-hat SEO and black-hat SEO.

Keywords: *Google Search, Page rank, SEO, SEO-Tools, Web-traffic*

I. INTRODUCTION

This section analyzes the work of previous research done on the field of SEO. To get a better understanding of the critical role SEO plays in web search, a quick overview will be given on each of the studies, highlighting the importance of the research.

II. LITERATURE SURVEY

2.1 Study and Analysis of Key Influence Web Search Factors

The study in [3] used a reverse engineering approach to “study and analyze the key influence factors in the process of web search.” Using this methodology, the researchers determined top five factors for SEO. The researchers developed a system that crawled all website factors (e.g. HTML structure, URL length, etc...) for 200,000 web pages using 10,000 search keywords as their sample set. Moreover, the keywords in the sample set were divided into the following three segments according to their Google search volume in the past three months: **Hot** – high search volume, **Middle** – medium search volume, **Cold** – low search volume.

The objective of this categorization was to discover the “different SEO factors on different segments” [3] of keywords. That is, the researchers were looking to uncover if keywords with different search volumes (i.e. low search volume, medium search volume and high search volume) required different SEO strategies; it may be case that more competitive keywords (the *Hot segment*), the ones with the highest search volume, required a different SEO approach than the less competitive keywords (the *Cold segment*).

Using an empirical approach to develop a series of analysis, their study determined top five factors for SEO that have the greatest influence in the natural or organic search for increasing high search rankings. Based on their research, they obtained certain SEO rules and provided valuable guidelines for SEO engineers to help improve website rankings. Table-1 lists the top five factors from their research survey.

Table-1: Top 5 SEO factors in study [3].

Rank	Hot	Middle	Cold
1	URL Length	URL Length	URL Length
2	Keyword appear in URL domain	Keyword appear in URL domain	Keyword appear in URL domain
3	Keyword density in H1	Keyword density in URL path	Keyword density in TITLE
4	Keyword density in TITLE	Keyword density in TITLE	Keyword density in TITLE
5	URL Layers	URL Layers	Keyword density in H2

As you can see from Table 2.1, the study showed *URL length* as the most important factor. Another important factor was the importance of placing the keyword within important web page elements: within the *URL*, *heading tags*, *domain name* and *title tag*. This means that the content of the web page must be relevant to the keyword. The difference between the SEO factors, among the three segments, may be due to the fact that high search volume keywords are more competitive; thus, they require a whole different SEO strategy. Surprisingly, one thing that was lacking from the paper was that the researchers didn't explain their definition of *URL layers*. Furthermore, *keyword density* refers to the number of times the keyword appears throughout the HTML page, HTML tag or SEO factor such as the URL.

In conclusion, this paper offers valuable insight and suggestions for SEO engineer's to follow when optimizing websites; and if implemented correctly, high rankings for specific keywords can be achieved. Author has discussed three SEO techniques were also implemented: *keyword appearing in the site domain*, *keyword appearing in the H1 tag*, and *keyword in the HTML title tag*.

2.2 An Empirical Study on the SEO Technique and Its Outcomes

This study is based on use of the Chinese search engine Baidu; Google's *PageRank* algorithm is not considered. Instead, the authors defined a metric (*Page Interest*) and consider whether certain SEO methods have any influence on it. Even though this study doesn't apply to Google, I felt its relevancy was important to analyze similarities from selected SEO factors from different search engines. Using data collected from 116 websites, the researchers sought to “analyze the impact of SEO techniques” and determine “which technique strategy was more effective” [2]. The following metrics were selected to measure the effectiveness of their SEO methods. The researchers believed these metrics were positive indicators of the SEO implementation:

- Indexed pages: the number pages crawled by the search engine bot.
- Number of independent IP address (IP): the number of different IP addresses accessing the web site.

- Pageview (PV): a user request to load a single HTML file.
- Reach: the percent of global Internet users who visit the site.
- Page view per user (PV/U): the average number of pages viewed by the total number of visitors to the web site.

Furthermore, the researchers also tested any correlation of SEO techniques on *Page Interest*, an additional metric they derived, which will be explained next. *Page Interest* “indicates the interest users show on Page” [2]; this means that the higher the *Page Interest* the higher the preference users will show to a website. According to the authors, *Page Interest* is related to *Pageview* and *Bounce Rate*, which is “the percentage of websurfers who visit websites and quickly leave” [2]. In other words, it’s the percentage of pagevisitors who decide that the page is not relevant to their search query and quickly leave. Thus, a low *Bounce Rate* means users who visit a web page stay there longer because they found what they were looking for; the keyword used for searching was found to be highly relevant to the content of the web page. The authors defined a candidate metric, *Page Interest*, and then investigated the effect of several SEO techniques on this metric. Below is the formal definition of *Page Interest (I)*:

- $I = \text{URL.pv} \times \text{URL.time} / \text{URL.bounce}$
- **URL.pv**= the number of page views on average per day
- **URL.time**= the time users spend on the website (in minutes)
- **URL.bounce**= the bounce rate

From the above definition of *Page Interest*, one can see that if there is a high number of *Pageviews*, an increase of time spent on the website and a low *Bounce Rate*, there will be a high *Page Interest*. As stated previously, a high *Page Interest* can be seen as positive interest the users show on the website; in other words, the website in question is highly relevant to the user’s search query. Below were the six SEO techniques that the researchers implemented and analyzed in their research:

1. **Overall Links** – The total number of web pages linking to another website. This has been a huge determining factor on how search engines determine a site’s position in the search results.
2. **Website Title Length** – Most search engines use the title tag on the search results page. Search engines also use the title tag to determine the theme or what the webpage is about. Therefore, optimizing the title tag is important.
3. **Keyword Density** – This is defined as the percentage of times the keyword appears on the web page compared to the total number of words on the web page.
4. **Layer Number** – This is “related to the logical structure of a website designed according to the relationship between content relevance and link position” [2].
5. **Page Size** – The researchers defined it as “the sum of the file sizes for all the elements that make up a page” [2]. According to them, “most search engines will not fully index pages that are greater than a certain size” [2]; therefore, the smaller the web page size, the faster it will load.
6. **Customization of 404 Error Pages** – An Error 404 “Page not found” is displayed whenever a visitor requests a web page that no longer exists.

The final results of their study indicated that *Page Size*, *Customization of 404 Error Pages* and *Overall Links* are significant factors in the effectiveness of SEO. But as stated before, it’s important to note that this study was focused on Baidu, China’s most popular search engine. Although the research was performed on a different search engine,

Overall Links seemed to coincide with Google's ranking algorithm as an important factor [11]. Furthermore, *Pageview and Bounce Rate* metrics were also used in my study to measure the effectiveness of SEO on my experimental website.

2.3 SEO Research Based on Six Sigma Management

The authors of this paper conducted research and empirical analysis to determine what factors had the most positive effect on SEO and proposed a method for its implementation. The goal of the study was to help SEO engineers identify the most influential factors for SEO and how to manage the execution of these strategies by using Six Sigma Management model. Originally developed by Motorola in the 1980s as a business management strategy, Six Sigma "seeks to improve the quality of process outputs by identifying and removing the causes of defects (errors)" [11]; it was interesting to see that the researchers included this process model in their research as a way to execute their SEO methods more effectively. The top five SEO influence factors from their study can be seen in Table-2 below.

Table 2.2: Top 5 SEO influence factors in study [10].

Number	Influence Factors
1	The application of keywords in website TITLE
2	The application of keywords in web page content
3	The number of external links
4	The number of backward links
5	The number of paper indexed

Based on their Six Sigma approach and their research, the researchers propose a "basic flow of the website search engine optimization (SEO)" [10] process using the following strategic steps:

1. **Keyword selection and application** – their tests suggest that keyword selection is the "most important factor which influences the search results ranking" [10].
2. **Building external links** – According to their tests, they confirmed that "the number of external links, backward links and websites indexed have a positive correlation with the search results ranking" [10]. This idea is similar to Google's proprietary *PageRank* algorithm.
3. **Flow monitoring and search engine analysis** – The key to having a successful SEO implementation is to constantly monitor "the whole website and the flow of each page" [10]. Due to the frequent changes of search engine algorithms, it's important to constantly measure and analyze website data (Google Analytics) to verify any changes in the site rankings in order to "make instant adjustments to ensure rankings" [10].

The above suggestions for more effective SEO do correlate with my research. Although the research was done on Baidu and not on Google, it's important to know the similarities between different search engines and the critical page elements that have the greatest influence on search engine rankings. This knowledge will help SEO engineers understand the different algorithms and compare the most factors used in their ranking algorithm.

2.4 How to Improve Your Google Ranking: Myths and Reality

This study focused on the Google's ranking algorithm; the researchers sought to systematically "validate assumptions others have made about this popular ranking algorithm" [8] and identify what page factors or other criteria, had the most influence in its ranking algorithm. They designed and developed a ranking system to determine the most important factors Google uses to rank pages. Using a reverse engineering approach, the paper showed how their ranking system "can be used to reveal the relative importance of ranking factors in Google's ranking function" [8]. Thus, the paper provides guidelines for SEO engineers on what factors are the most critical for optimizing web pages in order to achieve higher rankings. Although it has been known that Google uses more than "200 factors" [10] in their search engine ranking algorithm, this study determined a subset of those factors. The researchers' top 5 SEO factors are listed below:

1. **PageRank** – how authoritative is the site (as determined by Google's algorithm)
2. **Domain** – the keyword appearing in the domain name
3. **Title Tag** – the keyword appearing in the title tag
4. **Description Tag** – the keyword appearing in the description tag
5. **URL** – the keyword appearing in the URL

PageRank was the dominant factor of what determines high page rankings. Most of the results of research paper [8] correlate with my findings and implementation of the SEO process. With the exception of *PageRank*, the SEO methods I used were focused on having the keyword within the *domain name*, *title tag*, *description tag* and *URL* of the page being optimized [4].

2.5 The Application of SEO for Internet Marketing: An Example of the Motel Websites

This study conducted an experiment similar to mine in that they used selected SEO techniques and applied them on a website over the course of a year and then analyzed the effects of the SEO. What was different from my study was that the authors implemented their techniques on an existing website; in my research, I developed a website (from the ground-up), applied selected SEO tactics and then measured the effectiveness of the SEO. The authors used an existing website *mymotel.com.tw* as a case study to apply the SEO techniques and selected *Janfusumas* their target keyword to optimize the site. The motel is located in southern Taiwan and it's in close proximity to the Jansufum World, a famous amusement park in Taiwan [9]. Table-3 below shows the research variables and their definitions that were used as metrics to measure and analyze the effectiveness of the SEO. In my research, *Number of Visits* and *Ranking* variables were also tracked and analyzed; *Pageviews* is another metric that was measured in my study.

Table-3: Research variables and operational definitions of study [9].

Research Variables	Operational Definitions
Number of visits	Total number of visit, Single visit to a website and then exit within a period of time. For example, it is counted as one time visit for reading all the information within an hour.
Pages	After deleting the unnecessary files, like picture, the total number of files provided by the server.
Hits	The total number of files provided by the server.
Bandwidth	Total of transmission
Ranking	After the user put keywords into the search engine, the website ranking order in the search engines shown on the result pages.

Below is a list of SEO strategies that were implemented on the existing website(*mymotel.com.tw*) for their target keyword (*Janfusum*):

1. Keyword was put in the HTML Title tag
2. Added the keyword to the ALT property of the image tags
3. Added keyword to the H1 header tags
4. Registered the website to the DMOZ open website catalog (DMOZ.org)
5. Directly submitted the website to the main search engines: Yahoo, Google and Bing
6. Executed WEB PING to the main search engines. Pinging notifies the search engines that the website has been updated; this increases the chance that these search engines will find and index the pages much faster.
7. Created profiles in popular discussion boards (forums) and put keywords into the signature lines (e.g. *Janfusum*)
8. Created a sub-domain with the keyword in it, <http://janfusum.mymotel.com.tw>
9. Created an XML sitemap for the search engines. Sitemaps lists all the pages in the website. Sitemaps tell search engines information about your site, how it's structured and how often to index (or crawl) certain pages.

According to their research results, the experimental website (*mymotel.com.tw*) moved higher in Google's search rankings results for the *Janfusum* keyword; the bandwidth also increased after applying SEO. The ranking for *mymotel.com.tw* went from the No. 14 position to No. 2 positions for their target keyword, and the bandwidth increased as a result of an increase of users who visited to the site. This study in effect shows the importance of SEO as a way to increase website rankings and traffic. The only back-linking strategy discussed in the study was the creation of profiles in forums and back-linking from the signature. While this may help in the rankings, there are more effective back-linking strategies that can be used as discussed in [4].

III. CONCLUSION

This Section briefly summarized previous research done on the field of SEO. As seen from their studies, effective SEO implementation can help websites attain visibility by achieving higher rankings in the search results. Although Google has stated that it uses more than 200 factors in their ranking algorithm [12], there really is no consensus on what the most important factors are that determine website rankings. As a result, SEO engineers have relied on their experience, or any research performed in this field, to determine the factors for a more effective SEO implementation. Two of the five papers ([2,10]) discussed the Baidu search engine and sought to identify what factors were important in its ranking algorithm. Although the results were not collected from Google, it's important to analyze any similarities from certain SEO factors between different search engines. In this case, hyperlinks were seen as a critical factor in both search engines for determining high website rankings, but in order to get a better understanding of hyperlinks and the importance search engines play in the current state of the World Wide Web, we need to discuss the history of the Internet and its beginnings.

REFERENCES

- [1] Lewandowski, Dirk: The Influence of Commercial Intent of Search Results on Their Perceived Relevance. iConference (ACM), February: 8-11, 2011.

- [2] Wang, et al :An Empirical Study on Search Engine Optimization Techniques and Its Outcomes. (IEEE), 2011.
- [3] Zhu, et al :Research and Anaysis of Search Engine Optimization Factors Based on Reverse Engineering. (IEEE), 2011.
- [4] Grzywaczewski, et al :E-Marketing Strategy for Businesses. (IEEE), 2010.
- [5] Neshat, HamedSadeghi: Ranking of New Sponsored Online Ads. (IEEE), 2011.
- [6] <http://trends.builtwith.com/analytics/Google-Analytics> (Date retrieved: 2/16/12)
- [7] <http://support.google.com/googleanalytics/bin/answer.py?hl=en&answer=99118> (Date retrieved: 2/16/12)
- [8] Su, et al :How to Improve Your Google Ranking: Myths and Reality. (IEEE), 2010.
- [9] Li-Hsing, et al :The Application of Search Engine Optimization Techniques for Internet Marketing: An Example of the Motel Websites. (IEEE), 2010.
- [10] Zhang, et al :The Research on Search Engine Optimization Based on Six Sigma Management. (IEEE), 2011.
- [11] http://en.wikipedia.org/wiki/Six_Sigma (Date retrieved: 4/14/2012)
- [12] <http://www.seomoz.org/article/search-ranking-factors> (Date retrieved: 4/14/2012)