Search Engine Optimization (SEO) Techniques For Websites

Srinivasachar G[1], C R Rakshith Bhat[2]
[1]Prof., Dept of Computer Science and Engineering
[2]Dept of Computer Science and Engineering
[1, 2] Atria Institute of Technology, Bengaluru, Karnataka, India

Abstract- *With the growth of the Internet in the recent years, search engines such as Google, Bing, and Yahoo are becoming more and more crucial and reliable. The role of search engines is to index billions of web pages and display only the most relevant results for a given search query. When creating the website, many webmasters forget to take into consideration an essential factor, which is making the world aware of their website. Most of the times the main focus is set on making the website as user-friendly as possible, stable, fast, and secure. However, all of these techniques can be useless if the website does not have any visitor or people just cannot find it. To solve this problem, and to enhance the structure of the website to become more search-engine-friendly, a web application is developed to analyse any given webpage and to provide information on how to enhance and improve its structure to make it more search-engine-friendly and to improve its ranking on search engines. This process is also known as Search Engine Optimization (SEO).*

CCS Concepts- Information systems → Web indexing

Keywords- Search engine; website; Search Engine Optimization.

I. INTRODUCTION

A search engine is an application used to help find information on the internet. A search engine is usually accessed through a website that allows users to search the content of billions of websites available online by entering a set of keywords into a search textbox [1]. The keywords submitted by the user will then be matched against the cached or indexed pages in the search engine's database and a list of references that match the keywords or the keywords' criteria will be returned [2]. Search engines index new web pages and update existing ones regularly to provide a quicker and more efficient search.

Search engines are not only used to search for text content. They can be used to search for images, videos, applications, books and much more. In addition, some search engines such as Google, which is the leading search engine

today, allows the search to be filtered by time, country, and city. There are two main types of search indexes on the internet: directories and crawler-based search engines. Unlike search engines which use sophisticated algorithms to locate and index sites, directories are maintained by human editors. Directories are usually made of a categorized list of links to other websites. Those websites are submitted to the directory manually and then reviewed by editors for approval and inclusion in the directory. Crawler-based search engines are different from directories in that they are not maintained by humans. They use smart algorithms called spiders or robots to search and index web pages. These spiders are always running, crawling around the web, searching for pages, and taking snapshots of those web pages to be indexed or stored in the search engine's databases. They are so sophisticated that they can go after links from one page to another and from one site to another. Google and Bing [2] are good examples of crawler-based search engines, while Yahoo is a hybrid search engine that consists of both directories and crawler-based search engine.

II. RELATED WORK

Different types of solution were developed, but few are practical enough to enhance the search engine optimization of a website or provide helpful tips on how to improve the ranking of a website. Several solutions exist on the market: "Webmaster tools", "Google Analytics", "Google PageSpeed Tool", "Xenu Link Sleuth", "Google AdWords Keyword Tool", "Raven Tools", and "Alexa".

Webmaster tools such as [3, 4] Google Webmaster Tools and Bing Webmaster Tools are free tools that give an important insight about a website. The tools allow a webmaster to see what search queries led to the site being listed in the search engine results pages, the number of impressions (how many users have seen the website in the results), the number of clicks, the click through rates, and the average position in the results. In addition, the webmaster tools provide a list of all the internal links, e.g. links within the website pointing to other pages in the same website, navigation button and menus, list of broken links such as links

to non-existing pages, and list of websites pointing to our website. Furthermore, webmaster tools allow a webmaster to submit a sitemap which contains a list of all the pages of the website and their links in order for the search engine to be able to index all the web pages. Sitemaps also contains the date when the page was last modified, the priority of the page and how often the page is updated. This tells search engines how often they should index that specific page and update their databases.

Google Analytics [5, 6] is a service offered by Google. It provides webmasters and marketers details about the visitors of the website, such as the number of visitors during a time frame, the visitor's location, time of the visit, average visit duration, new visitor or returning visitor and bounce rates.

In addition, it can provide information about the referring sites, that is from which website the user has come from. The referring sites can be any website, such as Google, Facebook, Twitter, blogs, and forums.

Google PageSpeed [7, 8] tools provide some experimental rules. These suggestions do not affect the overall PageSpeed score but can be useful to experiment new areas that might improve the overall speed of the website. Google PageSpeed tools are very effective and can be helpful as fast and optimized pages lead to higher visitor engagement, retention, and conversion.

Xenu Link Sleuth [9, 10] is a great free tool that crawls any website in a similar way to search engines and quickly finds broken links and other problems. Xenu Link Sleuth follows all the links on the website and eventually checks all the pages for broken links. Broken links can decrease the search engine rankings since there will be a lot of "page not found" errors. In addition, it provides details and overview of the website's pages and provides suggestions on how to fix and improve them. Xenu Link Sleuth is very useful specially on new websites because it can help in discovering the simplest mistakes to the most complicated mistakes. Xenu Link Sleuth [9, 10] can be downloaded for free and can be installed on any machine running a Microsoft Windows operating system. The verification of the link is done by searching for <a> HTML tags, images tags, and frames. It continuously updates the list of URLs until no new links are found.

Google AdWords tool allows one to find related keywords to the ones they enter and see how many searches all those keywords are receiving per month [11, 12, 13]. The results returned by Google AdWords Keyword tool contain

information such as the competition on the keyword, the global monthly search of this keyword, and the local monthly search.

Raven tools [14] take all the data that can be used in an online marketing campaign and bring it all together in one place. Raven Tools is an all in one SEO tool. The list of included tools are so long that it makes it hard to get familiar to it quickly. Raven tools include research tools to search for keywords and the competition of those keywords and the monthly search, in addition they include contact managers, keyword managers, use data from the Google AdWords campaigns such as reports and insights, and grab Google Analytics data from the account.

The Alexa traffic rank [14, 15] is taken from data collected from millions of users over a period covering the last three months. The collected information is based on the number of users who visited a specific website and the number of pages that were viewed by that same user. The traffic rank is updated almost every day depending on the page views and website visits, the lower the Alexa number rank is the higher the amount of traffic a website is receiving, for example a website ranked in position 1 would indicate that this website have the highest rank that any website can reach.

The traffic rank is based on the collected historical data from millions of installed Alexa Toolbars on user's machines. In addition, the data is also obtained from other traffic sources. The traffic rank is a combination of page views and number of visitors.

The initial step consists of a complex algorithm that computes the number of visitors and the number of page views for all the websites on the World Wide Web on a daily basis [15]. The Alexa traffic rank is derived from the results of this calculation averaged over time, so that the site rank is eventually reflecting both the number of visitors and the number of users who visited this website.

III. PROPOSED SOLUTION

The aim of our proposed solution is to produce a simple website that is both user-friendly and free, with the goal of testing the SEO on any given website and providing explanations on how to enhance it, taking into consideration the accomplishments and design of the existing solutions.

The proposed web application works on any modern browser, it scans the source code of any given webpage and displays results and explanations about the structure of that page, its weaknesses and strengths, how it can be improved

and made search-engine friendly. There are many aspects that affect the search engine optimization, from the essential structure that is done on-site to the outside factors that can affect the ranking of a website. In order to understand how our proposed solution works, one must first understand how the structure of a website should be built, how it affects the search engine optimization, how it should be used and what are the limits. In order to achieve this, detailed explanation will be provided about On-page SEO, Off-page SEO and the impact of Social Media.

ON-PAGE SEO

On-page SEO is a technique applied web pages to allow search engines to understand the content of a webpage. On-page SEO helps by optimizing the page structure for both search engines and users. The crawlers of the search engine will be able to read, analyse and index the content of the webpage easily and to provide better results to the users after they submit the search query. Many factors affect the on-page SEO techniques such as page title, header tags, Meta keywords and description, headers, ALT tags, URL structure and size, internal linking, sitemaps, webpage compression and robot.txt. These factors have a huge impact on the on-page SEO and will make a website both user-friendly and search engine-friendly. In the following subsections, a detailed explanation will be given for every factor to explain how they are implemented, how they affect the search engines and how they should be used. Understanding these factors will help designing and implementing our application which will test the structure of a website and provide details on how to enhance it.

OFF-PAGE SEO

Off-page SEO are the techniques used outside of the website content to increase the website's rank in search engines. Unlike On-page SEO techniques, the Off-page SEO is done offsite and is usually not visible to the website. Generally, off-page SEO strategies are applied after the website has been built and after the On-page SEO has completed. Applying Off-page SEO techniques is done gradually over time and plays an important role in improving the crawlability of websites. It consists of building backlinks for the websites, submitting the websites to online directories, increasing the number of visitors through blogs and online articles, and increasing the websites' reputation and visitors through forums, communities, and social media.

SOCIAL MEDIA

Many webmasters concentrate on building inbound links to increase their search engine rankings while overlooking the impact of social media. Social media is a very important aspect of search engine optimization (SEO), because it helps driving more targeted visitors to the website. The more visitors a website is having, the higher rank it will have in search engines.

Search engines such as Google are lately turning to social media such as Facebook shares and likes, Twitter posts and Pinterest pictures for insight on rankings. A strong social media presence is very important and vital for any website. Social Media Optimization (SMO) is not a onetime process, it's an ongoing campaign that's always driving high quality and targeted audience.

IV. DESIGN AND IMPLEMENTATION

In this section, we present the design specifications of our solution. We will start by explaining the Data Flow Diagram.

4.1 DATA FLOW DIAGRAM

The Data Flow Diagram shows in a top-down approach by the different checkpoints of our system as shown in Figure 1.
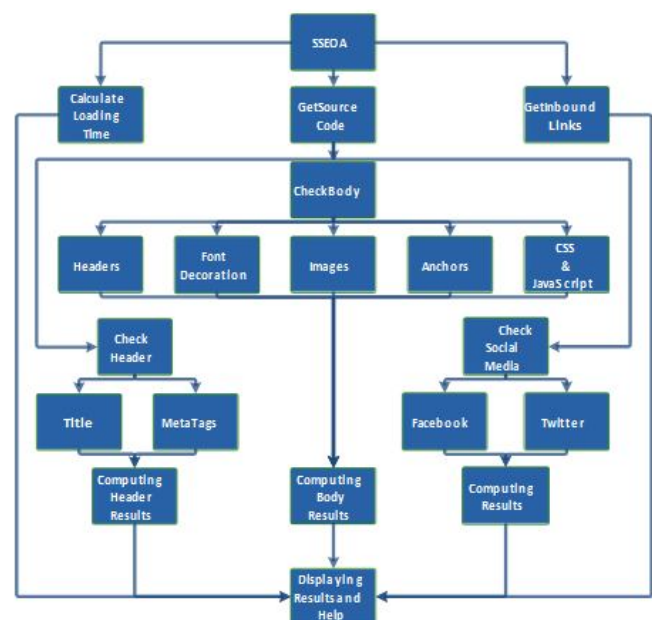

Figure 1. Data Flow Diagram

4.2 IMPLEMENTATION

The programming languages and Database Management System used to develop the web application are PHP, MySQL, and jQuery.

The PHP language is an open source server-side scripting language that is totally free and widely used. PHP files have by default ".php" as a file extension, and can contain text, CSS, HTML and PHP code. Once the PHP code is executed on the server side, the HTML is returned to the user's browser.

In addition, PHP is capable of doing everything that other languages can do, such as creating classes, encrypting data, creating forms, generating dynamic content, connecting to database and reading and writing cookies.

MySQL is a Database Management System used mostly on the web and runs on a server. It is widely used because it can operate on both small and large applications. MySQL is stable, fast, efficient, reliable, it supports all the standard SQL language, and is very easy to use. Furthermore, MySQL is free to download and install. MySQL is currently developed, supported and distributed by Oracle, and is compatible with all operating systems.

jQuery is a lightweight JavaScript library and has the role of making the use of JavaScript much easier for developers, as it can take several common tasks that usually require hundreds of lines in JavaScript and wrap them into one method that can be called with a single line of code. jQuery can also simplify the complication of AJAX calls and DOM manipulation, it also allows the manipulation of cascading style sheet (CSS) and animations.

The web application will be called "Simplified SEO Analyzer", the user will enter any given URL in the text box and click analyse as shown in Figure 2.

Figure 2. Simplified SEO Analyzer application

The above figure shows the application main page, the user will enter any web page URL into the text field, and when the user clicks the "Analyse" button, the application will fetch the HTML source code of the web page, and will analyse all its HTML elements.

The results will appear based on the missing or non-optimized elements as shown in the figure.

In the first "Results" tab, advises and tips will be displayed. The red colour indicates that the error is critical and it should be fixed for a better optimization. The orange colour indicates that the error is less critical but still important to be implement. The yellow indicates that the error is less important but can improve ranking.
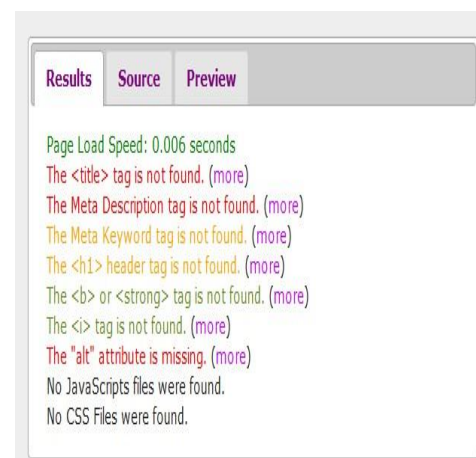.

Figure 3. Obtained Results

The light green indicates that the problem can slightly improve the performance.

The "Source" tab displays the fetched source code of the targeted website. This is useful for a dynamic website that has a lot of jQuery or AJAX calls. In addition, it helps displaying the source code that search engines are crawling and analysing.

The third tab "Preview", displays in an iframe the page being analysed, in order to visually make sure that the intended page to be analysed is actually being analysed.

V. RESULTS AND VALIDATION

First, the two websites were created, seo-testing.com and seotesting.tk, the web pages to be tested were uploaded to the server. Google, Bing, and Yahoo webmaster tools were used to manually submit the web pages to the search engine in order to get them indexed, to track the number of indexed web

pages and to submit a sitemap to the search engines to quicken the indexing process.

In order to produce the most accurate results, every experiment was conducted ten times using different web pages' names and content.

Both websites were uploaded to the server of each domain name. The servers are hosted in different locations with different IP addresses. The first experiment is to test whether search engines will index the web-pages without any backlinks and without any tool.

After one week of waiting time, Google was able to index two web pages out of ten for the .com domain and zero web page of the .tk domain. Bing and Yahoo did not index any webpage throughout the week period as shown in Figure 4.
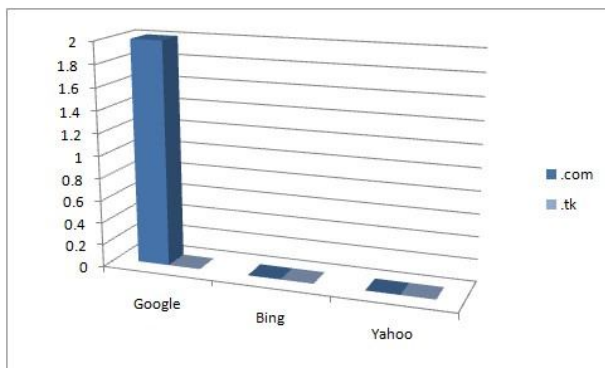

Figure 4. .com vs .tk – Week 1

During the second week, the domain names were added to Google and Bing webmaster tools, an XML sitemap was submitted, and the web pages were manually submitted to the webmaster tools to be indexed.

After the second week was over, and although, both domain names and all the web pages were manually submitted to be indexed, it seems that Google and Bing did not honourall these requests, and only few web pages were indexed. For the .com domain, six web pages were indexed on Google, and 3 web pages were indexed on both Bing and Yahoo. For the .tk domain, three web pages were indexed on Google, and 1 web page were indexed on both Bing and Yahoo.

The below graph shows the difference between Google, Bing, and Yahoo, and the indexing after two weeks of the web pages.
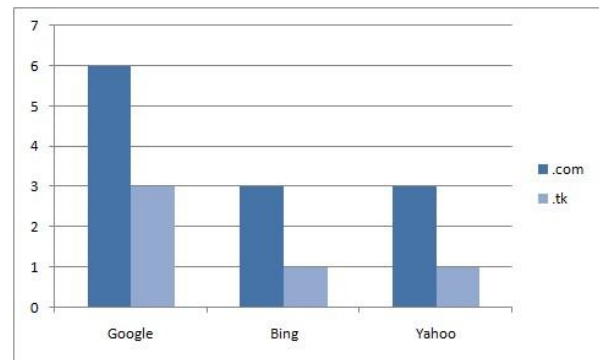

Figure 5. .com vs .tk – Week 2

Headers are usually used to divide a web page structure into sections and sub-sections and sub-sub-sections. Bold, italics and strong are useful for highlighting the keywords for the visitors in order to make it stand out from the other content.

The conducted experiment consisted of two identical fully optimized web pages. The only difference between the two pages is that one of them have headers, bolds and italics while the second did not.

The results showed once again that Bing discarded nine out of ten less optimized web pages, and only displayed the more optimized ones, and there was no way to check whether those less optimized pages were actually indexed or discarded.

On Google, the scenario was different, all the web pages appeared, and seven out of ten web pages with headers, bold and italicsranked above the ones without any headers. The below figure shows a comparison between Google and Bing.
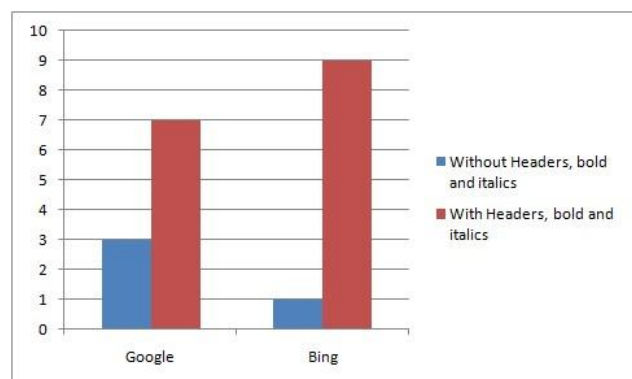

Figure 6. Comparison of Headers, Bold and Italics

VI. CONCLUSION

On-page SEO and off-page SEO are both techniques used to improve the visibility of a website. In addition to those two

techniques, social media is a very important aspect of search engine optimization (SEO), because it helps driving more targeted visitors to the website. The more visitors a website is having, the higher rank it will have in search engines.

On-page SEO is a technique applied on web pages to allow search engines to understand the content of a webpage. On-page SEO helps by optimizing the page structure for both search engines and users. The crawlers of the search engine will be able to read, analyse and index the content of the webpage easily and provide better results to the users after they submit the search query. Many factors affect the on-page SEO techniques such as page title, header tags, Meta keywords and description, headers, ALT tags, URL structure and size, internal linking, sitemaps, webpage compression, webpage load speed and robot.txt. These factors have a huge impact on the on-page SEO and will make a website both user-friendly and search engine-friendly.

Unlike on-page SEO which is applied on the website itself, off page SEO are the techniques used outside of the content of the website to increase the website's rank in search engines. Off-Page SEO is done offsite and is usually not visible to the website. Generally, off-page SEO strategies are applied after the website has been built and after the on-page SEO has completed. Applying off-page SEO techniques is done gradually over time and plays an important role for improving the crawlability of websites. It consists of building backlinks (inbound links) for the websites, submitting the websites to online directories, increasing the number of visitors through blogs and online articles, and increasing the websites' reputation and visitors through forums, communities and social media.

By using our developed web-application, a webmaster can analyse and improve the structure of the website based on the provided explanations, in order to improve the search engine visibility and achieve a higher ranking.

VII. ACKNOWLEDGMENTS

REFERENCES

[1] Ho Li-Hsing, Huang Jui-Chen, Lu Meng-Huang and Ho Hui-Yi, "The application of search engine optimization for internet marketing: An example of the motel websites" in Computer and Automation Engineering (ICCAE), 2010, pp. 380-383.

[2] Wu Di, Luan Tian, Bai Yan, Wei Liyuan and Li Yanhui, "Study on SEO monitoring system based on keywords & links" in Computer Science and Information Technology (ICCSIT), 2010, pp. 450-453.

[3] Janet Driscoll Miller, "Using The Structured Data Report In Google Webmaster Tools", SearchInsider, 2012.

[4] Google, Google Webmaster Tools [computer software], http://www.google.com/webmasters/tools, accessed on 16/03/2013

[5] Google, "Analysis of Users' Behavior on Web 2.0 Social Network Sites: An Empirical Study" in Information Technology: New Generations (ITNG), 2010, pp. 454 - 459

[6] M. Chafkin, "Analyze this, says Google," Inc., vol. 28, Apr. 2006, pp. 30-30.

[7] Google, Google PageSpeed [computer software], 2013, https://developers.google.com/speed/pagespeed/insights, accessed on 19/03/2013.

[8] Mo Yunfeng, "A Study on Tactics for Corporate Website Development Aiming at Search Engine Optimization" in Education Technology and Computer Science (ETCS), 2010, pp. 673-675.

[9] Ian Harac, "Xenu's Link Sleuth Finds Broken Web Site Links and Images", PC World Magazine, 2009.

[10] Tilman Hausherr, Xenu Link Sleuth [computer software], 2013, http://home.snafu.de/tilman/xenulink.html, accessed on 21/03/2013.

[11] Di Di, Luan Tian, Bai Yan and Wei Liyuan "Study on SEO monitoring system based on keywords & links" in Computer Science and Information Technology (ICCSIT), 3rd IEEE International Conference, 2010, pp. 450-453.

[12] Huanwei Wu, "Search Engine Optimization of E-Commerce Websites" in Management and Service Science (MASS), International Conference, 2011, pp. 1-3.

[13] Mason Zachary, "WordRank: A Method for Finding SearchAd Keywords for Internet Merchants" in Internet and Web Applications and Services, 2007, pp. 12.

[14] Raven, Raven Tools [computer software], 2013, http://raventools.com, accessed on 24/03/2013.