

# A Survey on Search Engine Optimization, Its Techniques, Tools and Algorithms

<sup>1</sup>P.Dinesh, <sup>2</sup>S.SenthilMurugan

**Abstract**— In today's competitive market, SEO is more important than ever. Search engines serve millions of users per day looking for answers to their questions or for solutions to their problems. If you have a web site, blog or online store, SEO can help your business grow and meet the business objectives. Certainly, how the more traffic can be increased to your website. This survey paper provides specific description for Search engine Optimization and explains about how it effectively works on a website to bring page rank and increase traffic towards your website, SERP (Search Result Page on Search Engine). This paper also focuses on latest SEO techniques and its algorithms.

**Keywords**— Search Engine Marketing, Website Promotion, Search Ranking.

## 1 INTRODUCTION

Search engine optimization (SEO) is the act of modifying the content and structure of web pages to increase traffic. SEO affects the user experience of a website, and the performance of that site's pages in organic search results on the major search engines (Google, Bing, Yahoo, Ask, AOL, etc.). More than all the Search Engine Optimization can be taken as a gaming technique that requires increasing the page rank of the website and also the traffic towards it. i.e., similar to a game that you play in your smart phone or your personal computer, and patiently and interestingly you will be seated to increase your scoreboard, this technique also provides various techniques and tools to play for it that could increase the traffic and page rank for the website.

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 any 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.

The majority of web traffic is driven by the major commercial search engines, Google, Bing, and Yahoo! Although social media and other types of traffic can generate visits to your website, search engines are the primary method of navigation for most Internet users. Search engines are unique in that they provide targeted traffic where people are looking for what you offer. Search engines are the roadways that make this happen.

Search queries are the words that users type into the search box. This carries an extraordinary value. It has been shown that search engine traffic can play a very huge role in an organization's success. Targeted traffic to a website can provide publicity, revenue, and exposure like no other channel of marketing. Investing in SEO can have an

exceptional rate of return compared to other types of marketing and promotion. Through boosting up of rankings, the content will be placed where the searchers readily find it. Unlike other major algorithms and processes, the Search Engine Optimization is less technical but it requires at least three to four months after a process to reach its result or page rank.

## 2 LITERATURE SURVEY

Search engines are a kind of tools which are designed to search information on the Web. The search engine results are generally displayed in a vertical sequence often referred to as search results pages. Links / URLs available on those pages are referred as hits. Search engine basically works on steps like, Web crawling, Indexing, Searching, etc. They work by storing information about web pages in databases, which they retrieve from the web. A search engine takes advantage of the hyperlinks that connect Web sites on the Internet. A software program called a Web crawler automatically browses the Web in a systematic way and sends out inquiries that "crawl" from site to site.

Since crawler is a software program, it is given different instructions on different computers. For instance, WebCrawler, a program launched in 1994, was the first software to index entire web sites rather than just page titles. Search engine crawlers operate within different sets of instructions or parameters, such as search titles and first paragraphs only, or to search entire documents, including metadata.

The information the crawler software collects automatically put into an index, when a query is submitted to that search engine's index. Each search engine has its own index. Thus the index searched by Google is not the same as the one searched by Yahoo or MSN (Microsoft Network). However, majority (80%) of Internet users are hooked on to three search engines - Google, Yahoo and MSN - Search.

The aim of SEO (Search Engine Optimization) is to get higher position of links in organic listings. Set of techniques are used for going up to the top of search engine listings. SEO has conceptually expanded to include all likely ways of promoting web traffic. Mainly there are two approaches for listing results on screen. One, organic (natural way listing) and second, pay per click (paid listing). Organic

- <sup>1</sup>P.Dinesh, Second Year Master of Computer Applications in Priyadarshini Engineering College, Vaniyambadi, E-mail: dineshdb693@mail.com
- <sup>2</sup>S.SenthilMurugan, Assistant Professor Master of Computer Applications in Priyadarshini Engineering College, Vaniyambadi, E-mail: mailmurugan.78@mail.com

search result listings appear as search results without the payment of a special charge to the search engine provider. Pay per click strategy is to gain company's revenue. SEO is mainly achieved by the combination of 2 main factors mainly on page and off page factors.

### 3.1 The major elements of SEO are

#### 3.1.1 Keyword-rich text

According to design matters to the search engines one can access the keyword-rich text, and it matters to human visitors so they can easily find that keyword-rich text once they arrive at site.

#### 3.1.2 Site and page architecture

To get optimized results from search engine robust site and page architecture plays vital role.

#### 3.1.3 Link development

This is one of the most overlooked components of a successful SEO. Link development can be defined as collection of links for the site from other web sites which to improve search engine ranking of the web site. The development of SEO methods also tacitly recognises the initiatives and market position of Google within the search engine domain. In effect other search engines indexing algorithms have had to recognise the preferences and indexing priorities of Google as a de facto standard. The consistency of approach has occurred despite the corporate secretiveness of the actual Google indexing algorithm - which, in itself, is not patentable.

## 4 HOW PEOPLE INTERACT WITH SEARCH ENGINES

Search engines have two major functions: crawling and building an index, and providing search users with a ranked list of the websites they've determined are the most relevant. Links allow the search engines' automated robots, called "crawlers" or "spiders," to reach the many billions of interconnected documents on the web. Once the engines find these pages, they decipher the code from them and store selected pieces in massive databases, to be recalled later when needed for a search query. To accomplish the monumental task of holding billions of pages that can be accessed in a fraction of a second, the search engine companies have constructed datacenters all over the world. These monstrous storage facilities hold thousands of machines processing large quantities of information very quickly. When a person performs a search at any of the major engines, they demand results instantaneously; even a one- or two-second delay can cause dissatisfaction, so the engines work hard to provide answers as fast as possible. Search engines are answer machines. When a person performs an online search, the search engine scours its corpus of billions of documents and does two things: first, it returns only those results that are relevant or useful to the searcher's query; second, it ranks those results according to the popularity of the websites serving the information. It is both relevance and popularity that the process of SEO is meant to influence.

Popularity and relevance aren't determined manually. Instead, the engines employ mathematical equations (algorithms) to sort the wheat from the chaff (relevance), and then to rank the wheat in order of quality (popularity). These algorithms often comprise hundreds of variables. In the search marketing field, we refer to them as "ranking

factors." The complicated algorithms of search engines may seem impenetrable. Indeed, the engines themselves provide little insight into how to achieve better results or garner more traffic.

An important aspect of SEO is making your website easy for both users and search engine robots to understand. Although search engines have become increasingly sophisticated, they still can't see and understand a web page the same way a human can. SEO helps the engines figure out what each page is about, and how it may be useful for users.

To perform better in search engine listings, your most important content should be in HTML text format. Images, Flash files, Java applets, and other non-text content are often ignored or devalued by search engine crawlers, despite advances in crawling technology. The easiest way to ensure that the words and phrases you display to your visitors are visible to search engines is to place them in the HTML text on the page.

Just as search engines need to see content in order to list pages in their massive keyword-based indexes, they also need to see links in order to find the content in the first place. A crawlable link structure is the one that lets the crawlers browse the pathways of a website. It is vital to them finding all of the pages on a website. Hundreds of thousands of sites make the critical mistake of structuring their navigation in ways that search engines cannot access, hindering their ability to get pages listed in the search engines' indexes.



Fig. 1 Example for Crawling error

In the example above, Google's crawler has reached page A and sees links to pages B and E. However, even though C and D might be important pages on the site, the crawler has no way to reach them (or even know they exist). This is because no direct, crawlable links point pages C and D. As far as Google can see, they don't exist! Great content, good keyword targeting, and smart marketing won't make any difference if the crawlers can't reach your pages in the first place.

### 4.1 Keyword research

Keyword research is one of the most important, valuable, and high return activities in the search marketing field. Ranking for the right keywords can make or break your website. By researching your market's keyword demand, you can not only learn which terms and phrases to target with SEO, but also learn more about your customers as a whole. It's not always about getting visitors to your site, but about getting the right kind of visitors. The usefulness of

this intelligence cannot be overstated; with keyword research you can predict shifts in demand, respond to changing market conditions, and produce the products, services, and content that web searchers are actively seeking. In the history of marketing, there has never been such a low barrier to entry in understanding the motivations of consumers in virtually any niche.

There are quite a few great, thorough keyword research steps to add search keywords for a website. In general, it is suggested to follow this process:

- Create a seed list of starting terms
- Expand your list using keyword research tools
- Refine your list with competitive research

The idea behind this model is that different keywords represent different levels of interest and intent; some search terms are used by people close to a conversion, while others are used by people still in the early stages of research. Keywords are categorized by type and arranged by how effective they are at converting.

## 5 SEARCH ENGINE OPTIMIZATION TECHNIQUES

There are two techniques to be followed for Search Engine Optimization



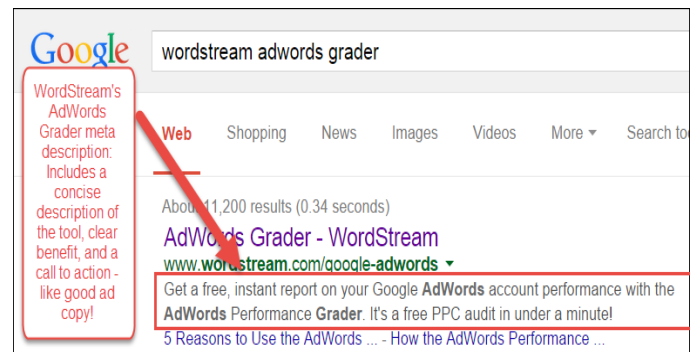
## 6 ON-PAGE OPTIMIZATION

On-page SEO is the practice of optimizing individual web pages in order to rank higher and earn more relevant traffic in search engines. On-page refers to both the content and HTML source code of a page that can be optimized, as opposed to off-page SEO which refers to links and other external signals. It refers to all the things that you can do directly on your website mainly on the source code or html code of the website. Following are some techniques or steps to be followed for on-page optimization

### 6.1 Meta tags Title and Description

While Google is working to better understand the actual meaning of a page and de-emphasizing (and even punishing) aggressive and manipulative use of keywords, including the term (and related terms) that you want to rank for in your pages is still valuable. And the single most impactful place you can put your keyword is your page's title tag. The title tag is not your page's primary headline. The headline you see on the page is typically an H1 (or possibly an H2) HTML element. The title tag is what you can see at the very top of your browser, and is populated by

your page's source code in a meta tag. While the title tag is effectively your search listing's headline, the meta description (another meta HTML element that can be updated in your site's code, but isn't seen on your actual page) is effectively your site's additional ad copy. Google takes some liberties with what they display in search results, so your meta description may not always show, but if you have a compelling description of your page that would make folks searching likely to click, you can greatly increase traffic. An example is given below:



## 6.2 Content Optimization

Content optimization is the process of optimizing web pages and their content to be easily discoverable by users searching for terms relevant to your website. The term SEO also describes the process of making web pages easier for search engine indexing software, known as "crawlers," to find, scan, and index.

## 6.3 Image optimization

The Alt tag is probably the most important image optimization factor. Alt stands for alternate. The text in the alt tag is displayed in place of the image if the image cannot be shown or is taking too long to load. Most of the major search engines are text-based, i.e. they cannot read images and videos.

## 6.4 Robot.txt

The robots exclusion standard, also known as the robots exclusion protocol or simply robots.txt, is a standard used by websites to communicate with web crawlers and other web robots. The standard specifies how to inform the web robot about which areas of the website should not be processed or scanned.

## 6.5 Xml site maps

Submitting the XML sitemap directly to their tools sites prompts them to crawl the sitemap more quickly, usually within the next few hours. In addition, the engines' tools sites provide additional data on the URLs in the XML sitemap, such as how many are indexed and whether the XML sitemap itself is valid.

## 6.6 Hatches files

The purpose of .htaccess files is to provide a means to configure Apache for users who cannot modify the main configuration file.

## 6.7 Canonicalization

For SEOs, canonicalization refers to individual web pages that can be loaded from multiple URLs. This is a problem because when multiple pages have the same content but different URLs, links that are intended to go to the same page get split up among multiple URLs.

## 6.8 W3C Validation

The Mark-up Validation Service is a validates by the World Wide Web Consortium (W3C) that allows Internet users to check HTML and XHTML documents for well-formed mark-up. Mark-up validation is an important step towards ensuring the technical quality of web pages.

## 7 OFF-PAGE OPTIMIZATION

Off-Page SEO refers to all the things that you can do directly off your website to help you rank higher, such as social networking, article submission, forum & blog marketing, etc. Following below are some techniques that are followed for off-page optimization:

### 7.1 Google Analytics

Google Analytics is a popular web analytics service that is offered by Google that tracks and reports website traffic. Google launched the service in November 2005 after acquiring Urchin. Google Analytics is now the most widely used web analytics service on the Internet.

### 7.2 Blogging

Blogging is one of the best ways to promote your website online! By writing a blog for your website, you give a reason for visitors to keep returning to your site and keep up to date with your latest posts. It also helps search engines to crawl your site more frequently, as they have to update your latest blog post entries, which ultimately helps you rank higher in search engine results pages (SERPs).

### 7.3 Search Engine Submission

Search engines will eventually find your site online, but that can take a while. To speed everything up, you should submit your website to the most popular search engines like Google, Yahoo, Bing, etc.

### 7.4 Directory Submission

Many people may say that directory submission is dead! But, it is not real. It is purely based on how effectively we are selecting those directories and how efficiently we are choosing the category for submission. You could submit to general directories, but for maximum effect, you are better off submitting to niche directories. Of course, I agree that it gives quite delayed results, but it is worth doing it.

### 7.5 Social bookmarking

Social Bookmarking is another great way of promoting your website. Submit your latest blog posts and pages to the most popular bookmarking sites, like Stumble Upon, Digg, Delicious, Reddit, etc. Search engines really like these types of sites because the content on these sites is updated very frequently. You should be very careful while doing this and you must properly handle the tags which are very essential to broadcast your news on a wide area network. This may increase your website traffic based on how effectively you have participated.

## 8 SEARCH ENGINE OPTIMIZATION ALGORITHMS

### Panda

Google Panda update imposed penalty on sites with low quality, spammy or thin content. Moreover, websites with Keyword Stuffing, Poor User Experience and plagiarism had to pay price.

Things to Avoid:

- (1) Duplicate Content (use Copy cape)
- (2) Plagiarism
- (3) Keyword Stuffing

## 8.1 Penguin

Google Penguin deals with link quality and penalizes those sites which purchase back links from third party sites. Also, Websites which don't follow Google Webmasters guidelines have to face the music.

### 8.2 Disavow Tool

Another main initiative of this update was the introduction of Google's Disavow Tool. With the help of this tool, we can remove any link which harms our site and inform Google not to consider link juice from that specific link.

### 8.3 Pirate

Google Pirate was launched to protect Intellectual Property. It penalized those sites which received copyright Infringement reports. Majority of sites which were affected were that made pirated content (music, movies etc).

### 8.4 Humming bird

Google Humming Bird was launched to manage "Long Tail Keywords" based upon user's intent. It takes into account search results that match searcher intent, rather than individual keywords within the query.

### 8.5 Pigeon

Google Pigeon Update provided more relevant, precise results based upon user's location and other geographical factors.

Due to this Local SEO became prominent and Google My Business became an essential ingredient which shows entities on map based upon three factors.

- (1). Relevance
- (2). Searcher Distance
- (3). Prominence

### 8.6 Mobile friendly

Google Mobile Friendly Update or also known as Mobilegeddon ensures that Websites which are mobile optimized <http://makeubig.com/2-017/09/12/amp-versus-responsive-web-design/> are given importance in SERPs (Search Engine Results Pages). So, Mobile Friendliness became an essential trait and Subsequently AMP (Accelerated Mobile Pages) was introduced to enhance speed.

### 8.7 Rank Brain

Rank Brain is a Machine Learning Artificial Intelligence algorithm which improved keyword search results based upon user's intent. Machine Learning is a process by which computer teaches how to do a specific task by itself, rather than being taught by human being. Through AL, Search engine acts smart as human beings by understanding real world problems and solving it. Using above principle, Rank Brain interprets user's search keywords that do not match the exact words they are searching for.

### 8.8 Fred

Fred targets websites that violate Google's webmaster guidelines. The majority of affected sites are blogs with low-quality posts that appear to be created mostly for the purpose of generating ad revenue.

## 9 EMD

The EMD (Exact Match Domain) is a filter Google launched in September 2012 to prevent poor quality sites from ranking well simply because they had words that match search terms in their domain names. When a fresh EMD Update happens, sites that have improved their content may regain good rankings. New sites with poor

content or those previously missed by EMD may get caught. In addition, "false positives" may get released

## 10 SOME COMMON REASONS WHY THE WEB PAGES MAY BE NOT REACHABLE

### 10.1 Submission required forms

If you require users to complete an online form before accessing certain content, chances are search engines will never see those protected pages. Forms can include a password-protected login or a full-blown survey. In either case, search crawlers generally will not attempt to submit forms, so any content or links that would be accessible via a form are invisible to the engines.

### 10.2 Links in un-parable JavaScript

If you use JavaScript for links, you may find that search engines either do not crawl or give very little weight to the links embedded within. Standard HTML links should replace JavaScript (or accompany it) on any page you'd like crawlers to crawl.

### 10.3 Links pointing to pages blocked by the Meta Robots, tag or robots.txt

The Meta Robots tag and the robots.txt file both allow a site owner to restrict crawler access to a page. Just be warned that many a webmaster has unintentionally used these directives as an attempt to block access by rogue bots, only to discover that search engines cease their crawl.

### 10.4 Frames or iframes

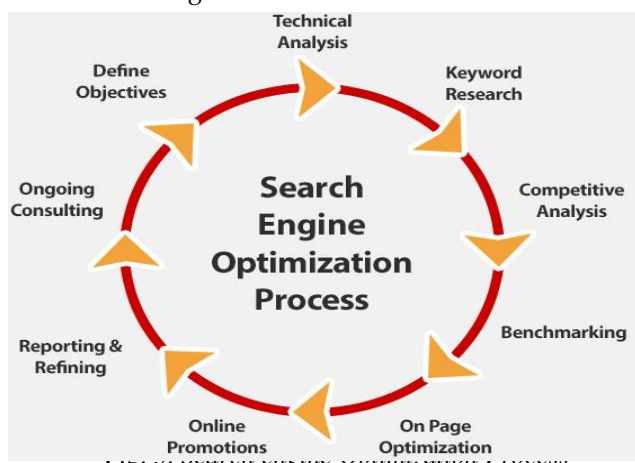
Technically, links in both frames and iframes are crawlable, but both present structural issues for the engines in terms of organization and following. Unless you're an advanced user with a good technical understanding of how search engines index and follow links in frames, it's best to stay away from them.

### 10.5 Robots don't use search forms

Although this relates directly to the above warning on forms, it's such a common problem that it bears mentioning. Some webmasters believe if they place a search box on their site, then engines will be able to find everything that visitors search for.

### 10.6 Links on pages with many hundreds or thousands of links

Search engines will only crawl so many links on a given page. This restriction is necessary to cut down on spam and conserve rankings.



## 11 MEASURING AND TRACKING SUCCESS

In search engine optimization, measurement is critical to

success. Professional SEOs track data about rankings, referrals, links, and more to help analyse their SEO strategy and create road maps for success. Every month, keep track of the contribution of each traffic source for your site, including:

### 11.1 Direct Navigation

Typed in traffic, bookmarks, email links without tracking codes, etc

### Referral Traffic

From links across the web or in trackable email, promotional, and branding campaign links.

### 11.2 Search Traffic

Queries that sent traffic from any major or minor web search engine.

Knowing both the percentage and exact numbers will help you identify weaknesses and give you a basis for comparison over time. Knowing the number of pages that receive search engine traffic is an essential metric for monitoring overall SEO performance. From this number, we can get a glimpse into indexation – the number of pages from our site the engines are keeping in their indexes. For most large websites (50,000+ pages), mere inclusion is essential to earning traffic, and this metric delivers a trackable number that's indicative of success or failure. As you work on issues like site architecture, link acquisition, XML sitemaps, and uniqueness of content and meta data, the trend line should rise, showing that more and more pages are earning their way into the engines' results. Pages receiving search traffic is, quite possibly, the best long tail metric around. While other analytics data points are of great importance, those mentioned above should be universally applied to get the maximum value from your SEO campaigns.

## 12 CONCLUSION

This paper demonstrates and explains the theory on Search Engine Optimization, how it works on today's marketing environment, how it can affect the traffic towards the websites, etc. It is defined and experienced that the search engine Optimization has a very important role in the development of the organization. And also this is the major technique that can be used in increasing the page rank and also increasing the data traffic for visiting the website. The more the visitors reaches or visits the website, the more the page rank increases. Search Engine Optimization is mainly a marketing technique that can play a major role by performing the on-page and off-page optimization. The directory submission, social bookmarking, etc will get into a very effective process for increasing traffic to the website. These techniques were very well demonstrated in the paper as an overview for Search Engine Optimization and its techniques. In future work to develop the one-page-optimization to avoid the unrelated links from first page its only depends by users for saving time.

## REFERENCES

1. Search engine Optimization with Google- IJCSI, Volume 9, Issue 1, No 3.
2. Search Engine Optimization-Wikipedia
3. Search engine Optimization-Search Engine Land, [www.searchengineland.com](http://www.searchengineland.com)
4. SEO Articles- Webconfs, [www.webconfs.com](http://www.webconfs.com)
5. Search Engine Journal, [www.searchenginejournal.com](http://www.searchenginejournal.com)

6. Search Engine Optimization: An hour a Day, Author: Gradiva Couzin and Jennifer Grappone
7. The Truth About Search Engine Optimization, Author: Rebecca Lieb
8. Singh T. and Maini R., "A Comprehensive Review on Search Engine Optimization",
9. Journal of Global Research in Computer Science, vol. 4, no.1, 2013.
10. Zona Market Bulletin , "The Need for Speed", , issue. All about SEO organic and inorganic SEO explained. Available at:
11. <http://www.tectutive.com/all-about-seo-organicand-inorganic-seo-explained/>.
12. Introduction to Website Speed Optimization, Available at: <https://kinsta.com/learn/pagespeed/>.

IJSER