A Study on SEO Monitoring System Based on Corporate Website Development

Rakesh Kumar[1] and Shiva Saini[2]

[1]Associate Professor, [2]Research Scholar, DCSA Deptt.,
Kurukshetra University, Kurukshetra,
rsagwal@gmail.com, shiva.saini@gmail.com,

Abstract: *Along with the fast growth of network information, using search engines to search information has developed into an important part of one's life everyday. In current years, there is a research focus on the search engine optimization technologies used to rapidly publish business information onto the search engines by which higher rankings can be reserved. The present paper analyzes the impact of receiving and recording of search engines and ranking rules to get understanding of the features of search engine algorithms frequently used and proposes the optimization strategy for the development of a website. This system provided good performance in monitoring the skills of SEO of website, and provided the consistent information support for ongoing optimization on search engine and making search engine marketing strategy. With the problems, the paper place ahead design Methods for website optimization, and summed up the optimization strategies for website design.*

Keywords: *Search Engine Optimization, Page Rank, Website Design Methods, Internet Marketing.*

1. Introduction

SEO (Search Engine Optimization) is the development of subsequent the principle of the search engine such as site structure, webpage language and interaction diplomatic strategies for the balanced planning to improve the site search performance in the search engine and increasing the opportunity of customer discovery and access to the website. SEO is a scientific development idea and methodology, which develops beside with the development of search engine, and promotes the development of search engines at the same time [15]. The major factors for good ranking positions in all the main search engines are:

1.1. On-page optimization −

Matching web page content and key phrases or keywords searched. The main factors to optimize web pages are keyword and key phrases density, keyword formatting, putting keywords in anchor text and the document meta-data (page title tags, Meta description tags, Meta keyword tags etc). Each web page on some website should have different title tag, Meta tag description and Meta tag keywords. The number of times the key phrase is repeated in the text of the web page is a key factor in determining the position for a key phrase [8].

1.2. External links (back links, backward links) and inbound links −

Google and other search engine counts link to a page from another website as vote for this page. In other words, web pages and websites with more external links from other websites will

be ranked more highly. The quality of external links is also very important factor, so if links from other web pages are from a website with a good reputation and with relevant content and key phrases, then this is more valuable. Many of web focused companies should improve this aspect of search engine optimization in process called external link building and internal link architecture. Web centered companies should always try to get links from web pages with the higher Page Rank (PR) or from web pages with a PR rating at least four. Page Rank is a scale between 0 and 10 used by Google to determine the importance of websites according to the number and quality of inbound links and external links (back links). Incoming or external links to website should always come in natural textual form, rather then in graphic form (Banner adds images etc). Google and other search engines does index or evaluate image links, but without textual links web companies won't increase website rank or popularity [20]. The Domain name maturity, Internal link structure, URL structure, Bounce rate etc. factors are very important for good ranking position in the search engine.

2. The Basic Concept of SEO

SEO, the most important method of network marketing, has turned into more popular in recent years. The main reason of SEO is to boost the experience of specific keywords with the purpose of increase website visibility, thereby growing sales opportunities. The major work of SEO is to examine how various search engines find the Internet page, how to index and how to conclude a particular keyword search results for ranking techniques. At the same time, we optimize the correlated web pages to get better ranking and it also improve site traffic, and enhance the site's technical capability which impact on advertising or promoting. The concept of SEO is initially proposed by Danny Sullivan, an industry analysis in 1997 [1] and since then it regularly is noticed by network operators and researchers in the search field. In 1998, Stanford University Sergey Brin and Lawrence Page proposed Page Rank algorithm, and at the same year, J. Kleinberg HITS algorithm is proposed while other scholars have proposed other link investigation algorithms, such as SALSA, PHITS, Bayesian and other algorithms [2]. Some of these algorithms have been implemented and used in the real search engine optimization system, and also we collect good results. Technically, SEO can be separated into White Hat SEO and Black Hat SEO [3]. White Hat SEO adopts the standard optimizing technology by search engine to have the regular and balanced operation; Black Hat SEO adopt malicious website ranking optimization techniques which violate the search engine optimization policy. The most important methods are as follows:

2.1. Doorway Pages

We use software to produce automatically a large number of pages with key words and then automatically revolve these pages to the home page and the reason is that these Doorway Pages with special key words find good search engine rankings.

2.2. Keyword Stuffing

Filling large quantity of keywords in the page, with adding the keyword density, we can enhance page significance for keywords.

2.3. Hidden Text

The keywords are placing into the HTML files. But, these words cannot be seen by users, and search engine can only distinguish.

2.4. Hidden Link

Place the key words into the link, and this link is indistinguishable to the user.

2.5. Cloaked Page

Programs or scripts are used to distinguish whether the access is the search engine or a normal user. If it is search engine, there will go back optimized web pages. If it is the regular people, the return will be another version [4].

3. Data Inaccuracy in Search Engine Optimization

These days, the total quantity of accessible information on Internet has become a danger. This type of condition when people are overfull with input result that people are unable to make any sense of so much data. One of the most determined problems on the Web is that managers can not precisely generate quality visitors to a website. As a result, websites are very frequently unused and fixed entities. In some cases, data inaccuracy is unseen killer of all web-analysis initiatives. Unfortunately, many of companies are inexperienced with collecting and interpreting data about website visitors. Certain companies buy the most expensive tool for data analysis. Then, they engage a well-respected web analysis consultant which measure and compare the wrong parameters. At the most companies, the web examination process is very effortless: they have some online tool which gathers online data. The data is collated and reported by a web analyst. Web analyst, at the most of cases, is not a qualified person for that kind of job. But for this kind of analysis web-driven companies must connect professional agency or person. Several issues and problems can be reason for data inaccuracy [5]. The important reasons of data inaccuracy follow:

3.1. Incomplete or incorrect tagging on a web pages –

Web tags are very important for tracking tools. When tags are not properly implemented on a site, a tracking tool can give imprecise data. In that case, tracking tool miscount website traffic and make reports with incorrect page names and confused paths.

3.2. Internal website visitors – People within companies use their website to look up information, assess its design etc. These visitors are not online customers and they should be automatically disqualified from analysis.

3.3. Complexity of search engine ranking factors – Google has acknowledged that it uses more then 200 factors or signals within its search ranking algorithms (Chaffey, 2010). Web analyst and web creators must have essential knowledge about these factors. In some cases, data inaccuracy might come as results of lack of information about search engine ranking factors.

3.4. Lack of search engine predictability – Some-times, search engines are very unreliable in terms of the website positioning. Website content creators and editors must be trained to understand the vital factors for the search engine ranking. Perhaps the main mistake companies make is presumptuous that good position of website must always be result of positive search engine optimization (SEO)[6]

3.5. Inaccurate cookies – Analytics software tools are mainly based on use of cookies in order to track users. These entirely harmless code which supply information to the analytics tool who is visiting the website, what they are doing and if they are new or repeat users. If the

analytics tool fails to place information on users PC or if online users delete cookies at wrong time, problems will take place.

3.6. Caching servers – Large companies occasionally use caching servers which automatically saves a local copy of some web page when online user visits a website. Companies organize these servers in order to cut down the amount of bandwidth customer's use [8]. In some cases, if online user decides to visit the same web page, the caching server deliver the stored page, rather then pulling down another from Internet. As result, web analytics tool might provide deceptive information about web traffic. Web analyst may conclude that their website is not performing well.

Problems with data accuracy may possibly lead to the following:

- *Missing website visitors* – The tracking tool does not record a real visitor data.

- *Tracking false website visits* – Tracking tool could report visitors who never visit a web page or site.

- Gathering of deceptive information for high volume metrics (visits, unique visitors, page views and other)

- *Poor decision making* – From a strategic point of view, data inaccuracy can reason that website owners decide to make needless changes an area of the website that is performing well.

- *Lack of faith in web analytics* – Website managers could challenge the roll of web analytics [9]. Over time, managers will start to question every data and therefore they will ignore the power of web analysis.

Data inaccuracy can lead to incorrect business decisions. Companies should always be attentive about accuracy of their data, because quality gathered and analyzed data are key factor for generating quality online visitors.

4. SEO STRATEGY

A number of websites rank good in the search engine, though several websites can't be found by the search engine, why? How to create the website has a high rank? Firstly we must to know the basic optimization strategy.

4.1. Page Title

Due to its maximum exposure in the page of search engine results, the page title is the main priority in the search engines, which is frequently taken as the first line of search results. Concerning the website click rate, the page title must be brief, eye-catching, and in order to attain the finest transformation, the page title should also be correctly summed up the contents of the page [15]. There are numerous attentions when set the page title:

4.1.1. It's better to build a concise introduction for the contents of the current page, and sequentially to allow the news expressed as clearly as possible, we should place significant keywords on the starting of the title.

4.1.2. Target keywords must be integrated in the html page title which is essential factor in the rank algorithm. Though too much repetition of keywords in the title is consider as awful strategy.

4.1.3. Set special title for all page from the particular visual angle.

4.1.4. Do not duplicate website navigation information in the title [11].
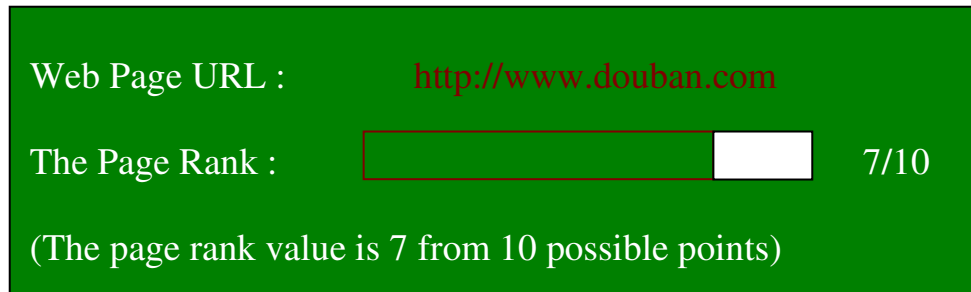

Figure1. The PR value of douban [10]

4.2. Keywords Optimization

Suitable keywords are the primary step to make a high-rank website. The higher the compactness of keywords in webpage, the closer significance between the webpage and the keywords, and the more forward place the page ranked in the search results. Keywords compactness is the ratio between keywords number and webpage words number, which are the mainly vital factors in optimization Strategy. In order to build website rank top in search engines, webpage must contain keywords as much as possible, but once keywords are extreme, it might face the possibility of registration cancellation. The prettiest compactness can be established with the help of first numerous highest rank websites in search results, for now keywords analysis tools must be used to analyze the position of keywords in search results for modification in time [11].

4.3. Improve the Click Popularity

Click popularity is a very significant factor for website rank in search engines. The popularity of the page frequently Clicked will be high. When a visitor clicks website from the search results, search engine will provide some certain value. But do not try to replicate clicking on your own website, because the clicks from a single IP will be considered only once [12].

4.4. Improve Link Popularity

Page Rank (PR) value is important for Google to judge the importance of website. If the linked page has a higher PR value, then the page has a higher PR value. Raising the value of PR can efficiently improve the website rank in search engine. Search engine calculate PR value through external website links quantities and the quality of links to Web sites, that is, it wants to improve the extensive degree of site's Link. This purpose can be achieved by exchanging links (between sites by adding links to each other). But it does not mean that can be to exchange links with any website randomly. It's better to prefer the PR value which is more than 4 [13]. Fig. 1 shows the PR value of douban queried by tool. We can notice that douban is a good website.

4.5. Internal Links of Website

Internal links of website pass on to the one another links between the different pages of website. Sensible links can build the web pages more visual, help visitors focus on the topic of website and make the topic key words more sensitive to the search engine. In addition, internal links of website can also increase speed the rate of the spider crawling [14].

5. Search Engine Algorithms

5.1. Page Rank Algorithm

Page Rank Algorithm, originate from Google's founder, Larry Page, is used to recognize the significance of web-pages separated in levels of 1-10 of which 10 represents full score, the higher PR value (Page Rank value), the more accepted the pages as in (1): when the PR value of a website has its score as 1, it means that this website has not popularity while the PR value of 7-10 shows that such website is very popular (or very important) [17]. In assumption Page-A has its page-t1, t2, and... tn linked to it, then Page-A has its PR value algorithm as follows:

$$PR(A)=(1-d)+d\{PR(t1)/C(t1) \ldots PR(tn)/C(tn)\} \ [16]$$

B. Hill Top Algorithm

There exists an essential fault for Page Rank Algorithm since a complete "importance value" is assigned to a webpage based on its linked quantity and quality. However, the "Page Rank value" is not a expression for a query, so even when a webpage just infrequently mentions key word departing from a thematic query in the content, the page can also get a higher ranking because of its high level of "Page Rank value" [18]. The Hilltop Algorithm is a speedy positioning method subject to classification of "experts" proposed by Krishna Bharat and George A. Mihaila, which has been applied by Google to the front-end of Page Rank. That is, the request is treated initially when it reaches using Hilltop Algorithm and given weight. Secondly, when nothing is searched out, it returns to "0" after which the Page Rank Algorithm is used for calculations. Finally, related ranking is conducted according to the result ahead the weighing of two algorithms.

C. New Algorithm for Search Engines

Currently, in order to provide users the correct results searched in a more scientific and balanced way, Google uses its sequencing technology combining Page Rank and Hill Top to establish the ranking of a webpage. Such a new algorithm has its formula as follows:

$$\{(1-d)+a(RS)\}*\{(1-e)+b(PR*fb)\}*\{(1-f)+c(LS)\} \ [17]$$

Of which, a, b and c are the regulation controls of weight and d, e, f are the damping controls while fb is the factor base. The new algorithm consists of three parts, the RS Correlation Scores, PR Page Rank Scores and the LS Industry [19].

6. Conclusion

This paper proposes websites design methods for SEO and introduces the essential SEO optimization strategy, which will facilitate to optimize the websites. The system design and

construction must be completely taken search engine marketing system development and the international situation in the future into account. For the need of the core work, we set up Search Engine Marketing network platform with the quality of high performance, high reliability, high accuracy and high security. Practices has proved that this system can give good performance in monitoring site in search engine optimization, and reliable information support for ongoing optimization on search engine of website, and making search engine marketing strategy. Though each search engine has its own algorithm & SEO strategies for website evaluation and ranking, fortunately there are common factors in the match between search terms entered and occurrence of the words on the web page. The number and the quality of external links are still recognized as the major ranking factor. However, with the development of search engine technologies, to discover continuously new tactics for optimization is clear to be a long-term and constant study.

REFERENCES

[1]. Wikipedia, "Search engine optimization," http://en.wikipedia.org/wiki/Search_engine_optimization Google, Google's Search Engine Optimization Starter Guide, 2008.

[2]. Fang Zhijian, Zhang Ruilin, Tong Xiaosu, "Recently research and future development of search engine," Computer Engineering and Design, Vol. 28, No. 16, pp. 4038-4041, 2007Burby, J., Atchison, S., Actionable Web Analytics, Sybex, Indianapolis, 2007.

[3]. Baidu, Inc. "Baidu Encyclopedia," http://baike.baidu.comlview/I047.htm?fr=alaO 1_1.

[4]. Wu Oil, Luan Tianl, Bai Yanl, Wei Liyuanl, Li Yanhui2, "Study on SEO Monitoring System Based on Keywords & Links", 978-1-4244-5540-9/10/$26.00 ©2010 IEEE.

[5]. Burby, J., Atchison, S., Actionable Web Analytics, Sybex, Indianapolis, 2007

[6]. Chaffey, D., Chadwick, F. E., Mayee, R., Johnston, K. Internet marketing, Prentice Hall, London, 2009.

[7]. SEOMoz. Google search engine ranking factors v2, published at www.seomoz.org

[8]. Google, Google's Search Engine Optimization Starter Guide, 2008.

[9]. Boris Knezevic, Mr.sc., "Search Engine Marketing As Key Factor For Generating Quality Online Visitors", MIPRO 2010, May 24-28, 2010, Opatija, Croatia.

[10]. "Check Page Rank of any web site pages," http://www.prchecker.info/check_page_rank.php

[11]. Jennifer Grappone,Gradiva.Couzin, "Search Engine Optimization," BeiJing: Tsinghua University Press, pp.100-147,July 2007.

[12]. "How to optimize Website?," http://blog.sina.com.cn/s/blog_4cedb9ba0100085g.html.

[13]. Dan Yi, "The exploration and realization of SEO," Computer Applications of Petroleum, vol. 14, No.1, 2006, pp.35-37.

[14]. Siyun Tan,Jiachao Zhu,Qianhua Zhou and Wei Xiong, "Research of SEO about Website," China Water Transport, vol. 08,Feb. 2008,pp.139-140.

[15]. Chengling Zhao, Jiaojiao Lu, Fengfeng Duan, "Application and Research of SEO in The Development of Web2.0 Site", preceding in 2009 Second International Symposium on Knowledge Acquisition and Modeling.

[16]. 2009 Research Report on Chinese Search Engine User Behavior (September 2009).(in chinese)

[17]. Liu Zhenghong, "Analysis on Google's New Algorithm and Tactics for User-tended Website Optimization", Inner Mongolia Science Technology & Economy. No.12, the 190th issue, Jun.2009. (in chinese).

[18]. Ju Jiehui, "Chinese Search Engines' PageRank Algorithm and Implementation". Computer Engineering and Design, 2007,28 (7): 1632-1635. (in chinese)

[19]. Mo Yunfeng, "A Study on Tactics for Corporate Website Development Aiming at Search Engine Optimization", 2010 Second International Workshop on Education Technology and Computer Science.

[20]. SEMPO, "The State of Search Engine Marketing, 2005", retrieved July 3, 2006, from http://www.sempo.org/learning_center/research/