SEO Techniques for various Applications - A Comparative Analyses and Evaluation

Sandhya Dahake[1], Dr. V. M. Thakare[2], Dr. Pradeep Butey[3]

[1](*sandhya.dahake@raisoni.net, G.H.R.I.T.*, Nagpur, *RTM Nagpur University, India)*
[2](*vilthakare@yahoo.co.in , P.G. Dept of Computer Science, SGB Amravati University, Amravati, India)*
[3](*buteypradeep@yahoo.co.in, Head, Dept of Computer Science, Kamla Nehru College, Nagpur, India)*

Abstract : *In this information times, the website management is becoming hot topic and search engine has more influence on website. So many website managers are trying o make efforts for Search Engine Optimization (SEO). In order to make search engine transfer information efficiently and accurately, to improve web search ranking different SEO techniques are used. The purpose of this paper is a comparative study of applications of SEO techniques used for increasing the page rank of undeserving websites. With the help of specific SEO techniques, SEO committed to optimize the website to increase the search engine ranking and web visits number and eventually to increase the ability of selling and publicity through the study of how search engines scrape internet page and index and confirm the ranking of the specific keyword. These technique combined page rank algorithm which helps to discover target page. Search Engine Optimization, as a kind of optimization techniques, which can promote website's ranking in the search engine, also will get more and more attention. Therefore, the continuous technology improvement on network is also one of the focuses of search engine optimization.*
Keywords*: Keyword, internet page, page rank, SEO techniques,*

I. INTRODUCTION

According to Wikipedia, "Search engine optimization (SEO) is a set of methods aimed at improving the ranking of a Web site in search engine listings…"

Search engine optimization, called SEO for short, is important to websites, which will improve the rank for search engines and get more page views. Search engine optimization, is a process of improving the prominence of a website. Search Engine Optimization is the practice of building a web site search engine friendly, so that it can be found easily on the search engine with its relevant keywords [1]. There are many free lancing SEO companies, which provide such facilities. The main role of these companies is to list the websites on search engine.

SEO is not a suddenly appeared technology, but it is synchronously developed with search engines. With the rapid development of information technology, SEO technology has attracted more and more attentions. In order to improve their website visit quantity, SEO techniques can make a better ranking in the search result using the keyword selection and deployment, high quality back links, rational website constitution, and rich content, etc. [2]

II. PREVIOUS BACKGROUND

2.1 THE CONCEPT OF SEARCH ENGINE

Search engine is a system is which based on certain strategies using specific computer program to gather information from the Internet. Search engine not only the necessary functions in the website to provide convenience to users, but also it is an effective tool of web user behavior. Efficient search engine allows user to find target information quickly and accurately[1]. The excellent search engine should have four characteristics: rapid, accurate, easy to use and strong. The information that search processes searched has high precision and meet requirement of user. In addition, easy to use is one of the reference standards in choosing search engine have the following two aspects. The one factor is whether the search engine can search the entire Internet, not just the World Wide Web. The other one is whether the search engine can change the length of the description or change the number of displaying the results page.

The ideal search engine should have both, a simple query capability and advanced search functions after search results come out[3]. According to difference of information collecting methods and service providing way the different types of search engines are available.

2.2 SEARCH ENGINE OPTIMIZATION

Search Engine Optimization, a popular network marketing method in recent years, focuses on increasing the web visibility through raising the exposure of specific keywords to create more sales opportunities. With the help of these techniques, SEO committed to optimize the website to increase the search engine ranking and web visits number and eventually to increase the ability of selling and publicity through the study of how search engines scrape internet page and index and confirm the ranking of the specific keyword. [3]

SEO, a guiding theory for search engine marketing, is not only about the ranking of search engine, but also about the every detail of web producing, construction, and maintenance, which worth the attention of every web producer, developer and agent to the significance of their job for the SEO.

2.3 WEBSITE SEARCH ENGINE OPTIMIZATION PROCESS

Search Engine Optimization is the practice of building a web site search engine friendly, so that it can be found easily on the search engine with its relevant keywords [1]. Keyword is most important factor which influences search result ranking; therefore the first thing is to solve the keyword selection and application. Building of external link, backward link and index have a positive correlation with search result ranking. Therefore, the superior building of linking and indexing after the perfection of keywords will greatly contributes to the SEO of the websites. Flow monitoring and search engine analysis. The key of SEO is the instant monitoring of the whole website and the flow of each page. The basic flow of website SEO as shown in following figure 1:
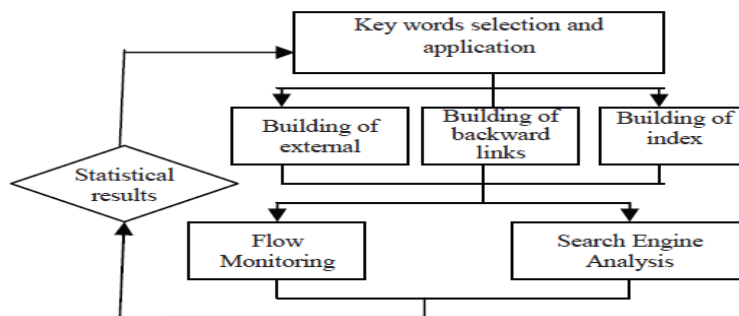
Figure 1: Web Site Search Engine Optimization Process

III.   VARIOUS SEO TECHNIQUES

There are many free lancing SEO companies [3], which provide the facility of building a web site search engine friendly, so that it can be found easily on search engine with relevant keywords. The main role of these companies is to list the websites on search engine. Some of those small SEO providers or other people also use some automated tools and/or other unethical techniques. The different SEO techniques are used for this purpose. Large web search engines use significant hardware and energy resources to process hundreds of millions of queries each day, and a lot of research has focused on how to improve query processing efficiency. One general class of optimizations called early termination techniques is used in all major engines, and essentially involves computing top results without an exhaustive traversal and scoring of all potentially relevant index entries.

SEO techniques also obtained various kind of identification from general website managers and even search engine inside [2]. But the arrival of foreign enterprise, our technology will gain upgrades. Large web search engines have to answer hundreds of millions of queries per day over tens of billions of documents. To process this workload, such engines use hundreds of thousands of machines distributed over multiple data centres. In fact, query processing is responsible for a significant part of the cost of operating a large search engine, and a lot of industrial and academic research has focused on decreasing this cost. Major families of techniques that have been studied include caching of full or partial results or index structures at various levels of the system, index compression techniques that decrease both index size and access costs, and early termination (or pruning) techniques that, using various shortcuts, try to identify the best or most promising results without an exhaustive evaluation of all candidates.

Comprehensive search engine to meet a lot of horizontal information search, the accuracy of search results, and the results of correlation and the user's search goal has some differences on the information needs of the relative concentration of the lack of classification more detailed industrial customers oriented[3]. With increasing the amount of information contained in search engine, in the mass of information queries to the information about the time it takes is gradually increased.

Search Engine Optimization Techniques are used for the following some of the applications [1-5]:
a) Spamming application
b) Block-Max Indexes application
c) Website application
d) News Industry application
e) Six Sigma Management application

IV. ANALYSIS AND DISCUSSION

a) Spamming application: Search engine spamming is a practice of misleading the search engine and increasing the page rank of undeserving websites [1]. This is a new way to counter those techniques using link based spam detection combined with the page rank algorithm. This technique helps us to discover target page and trace down the entire graph responsible for spreading spam.

The spam technique is mainly classified into two basic categories, namely; boosting technique and the hiding technique [1]. The boosting technique is the technique, which is used to make the page look more relevant to the search engine. The boosting techniques are further classified as keyword stuffing and link building. Keyword stuffing is also known as on page technique. During the on page, the target keyword is stuffed into the web pages i.e. the HTML page, PHP page or any other available source page on the web server. These keywords are stuffed into HTML tag i.e. META tag, H1 tag, HEADER tag etc. Each tag is rated explicitly by the search engine and the summation of all the ratings provides the total keyword density for the particular page. Even the sub directory, URLs and contents are rated and included in calculating the keyword density for the website.

The various factors affecting search engine results are: trusted link, link population, traffic, converging graphs. Hiding techniques are the techniques that are used to hide the boosting techniques [3]. These techniques are responsible for generating traffic from the user and misguiding the crawlers.

b) Block-Max Indexes application: Large web search engines use significant hardware and energy resources [2] to process hundreds of millions of queries each day, and a lot of research has focused on how to improve query processing efficiency. One general class of optimizations called early termination techniques is used in all major engines, and essentially involves computing top results without an exhaustive traversal and scoring of all potentially relevant index entries. There are several early termination algorithms for disjunctive top-k query processing, based on a new augmented index structure called Block-Max Index [2] that enables aggressive skipping in the index. Major families of techniques that have been studied include caching of full or partial results or index structures at various levels of the system, index compression techniques that decrease both index size and access costs, and early termination (or pruning) techniques that, using various shortcuts, try to identify the best or most promising results without an exhaustive evaluation of all candidates.

An augmented inverted index structure called a Block-Max Index and used it to speed up the WAND and Maxscore approaches [2]. The basic idea of the structure is very simple: Since inverted lists are often compressed in blocks of say 64 or 128 postings, we store for each block the maximum term score within the block, called block Maxscore. Thus each inverted list maintains a piece-wise constant upper bound approximation of the term scores.

c) Website application:
The search engine deals with tens of thousands of information search; the process is bound to follow the rule that pre-determined search engine operating principle. Any search engine after finishing one work will request in accordance with three following steps [3]:

1. **Crawl Page:** Each individual search engine has its own web capture progress. It along the hyper link of the web, continuously capture the pages. The capture page is called web page snapshot. Due to the application of hyperlink Internet is common, theoretically, starting from a range of web pages, you can collect the vast majority of page pages.

2. **Processing Page:** After catching web pages, still need to do lots of pre-treatment projects to provide retrieval service, among them, most important part is extracting keywords and establishing index file. Others also are including removing duplicate web pages, participles, judging the types, analyzing hyper links and counting pages important degree/abundance etc.

3. **Providing Search Services:** User inputs the keywords then Search engine finds the matching pages from indexed database; except for page title and URL, it still provide an abstract from web pages and other information to make users estimate expediently.

The SEO tool is tool that optimizes search engine function. An application of search engine optimization tool is designed to test that web site may obtained search ranking effect to search engine optimization degree machine. SEO tools mainly include keyword too, link tool, usability tool and other tools.

d) News Industry Application:

Existing news sites usually have the following two questions: (1) information update is slow. Slow information update is a common problem with the existing website. (2) News editor is poor. Existing news networks, including Sina-Chinese news web site, Yahoo China, Sohu and other influential online news media, there is the problem of poor news editor. In the case of limited human and material resources, it is hard to resolve. Then the news source for automated production will be reasonable and efficient, vertical search engine is suitable to the development of the news industry.

The vertical search engines [4] must meet the performance requirements of multi-threaded crawling. On this basis, it can be extending to distributed systems. The news site for news and information requirements is timely, accurate information. Lucene is an open source tool package. Its basic function is from the file system, database, Web, manual input and other sources of data integration, indexing documents, stored in the index database. When the user submits the query in the index quickly, it find the information in returned to the user. Lucene as a full-text search engine, it has the number of advantages.

e) Six Sigma Management Application:

SEO, a guiding theory for search engine marketing, is not only about the ranking of search engine[5], but also about the every detail of web producing, construction, and maintenance, which worth the attention of every web producer, developer and agent to the significance of their job for the SEO.

Six Sigma is a typical process management to conduct empirical study of SEO. Nowadays, Six Sigma Management is considered as the most efficient way to improve quality, reduce cost and raise effectiveness. Sigma (Σ,σ), originated from Greek, is a statistic unit used to measure the standard deviation among the total. While using Sigma to show the quality control level, 3σ control level means the product percent of pass is above 99.73%, while 6σ control level means the product percent of pass is above 99.99966%, and the defects per million opportunities (DPMO) is above 3.4 (or the Process Capability Index $Cp\geq2.0,Cpk\geq1.5$). Now 6σ are more than the statistical significance or a quality goal, more importantly, it has become an integration of philosophy, culture and method system.

Six Sigma Management includes DMAIC and DFSS, and DMAIC includes five phases, which mainly used methods are as listed in "Fig. 2".DFSS, short for Design for Six Sigma, is the design method for new process and new products.
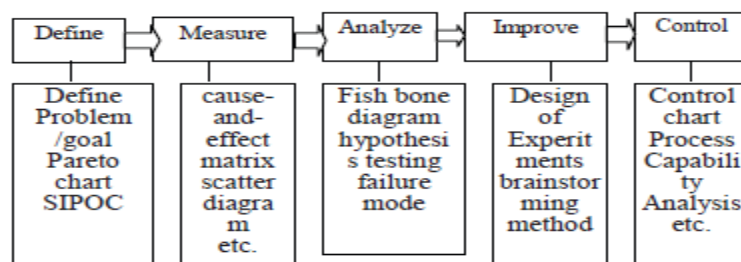

Define	Measure	Analyze	Improve	Control
Define Problem /goal Pareto chart SIPOC	cause-and-effect matrix scatter diagram etc.	Fish bone diagram hypothesis testing failure mode	Design of Experiments brainstorming method	Control chart Process Capability Analysis etc.

Figure 2: The Tools used by DMAIC Approaches in Six Sigma

V. CONCLUSION

In this paper, the comparative analysis and evaluation of different SEO techniques for various applications are included. These techniques have many advantages in web page designing. The importance of this paper is to suggest how we design a web site by using these techniques to get internet ranking on first page. According to the purpose of web page, we can use these techniques. Each technique has some importance, advantages as well as limitations.

References

[1].   Antriksha Somani, Ugrasen Suman , "Counter Measures against Evolving Search Engine Spamming Techniques", IEEE Conference, 2011 , 214-217.

[2].   Constantinos Dimopoulos, Sergey Nepomnyachiy, Torsten Suel, "Optimizing Top-k Document Retrieval Strategies for Block-Max Indexes", IEEE Conference, ACM, 2013, 113-122.

[3].   Meng Cui and Songyun Hu, "Search Engine Optimization Research for website Promotion," in *IEEE Int. Conf .of Information Technology, Computer Engineering and Management Sciences)*, 2011, 100-103.

[4].   M. Li, X.J. Gu, Z.X. Yang, "Research of Vertical Search Engine in News Industry," in *Proc. IEEE ISMOT*, 2012, 2012, 253–256.

[5].   Lihong Zhang, Jianwei Zhang, Yanbin Ju, in " The Research on Search Engine Optimization Based on Six Sigma Management", in IEEE conference 2011, 101-103.