OPEN ACCESS

International
Journal
Of Modern Engineering Research (IJMER)

An Innovative Approach to Web Optimization Based on User Profile for Online Searching

Mahendra Pratap Singh Dohare[1], Dr R. S. Jadon[2], Devendra Kumar[3]
[1]*Ph. D. Scholar, Dept- Computer Application MITS, Gwalior*
[2]*Prof & Hod, Dept- Computer Application MITS, Gwalior*
[3]*Assistant Prof Dept- Computer Science & Application Jiwaji University, Gwalior*

ABSTRACT: *Search Engine Optimization (SEO) has the vital to capitalize a business and also it is the one of the main key factor for marketing an online business. Dynamic websites are commonly used for e-commerce because they are easier to update and expand but they are subjected to indexing related problems. In this research paper aims to examine and address dynamic websites indexing related issues using efficient search engine optimization techniques. To achieve aims and objectives of this research we intend to explore dynamic websites indexing considerations, investigate SEO tools to carry SEO campaign in three major search engines, experiment SEO techniques, and determine to extent dynamic websites made search engine more efficient and friendly on these major search engines. We used SEO techniques which explored during this research to help dynamic webpage get indexed in major search engines. The experiment results reflect the effectiveness of SEO techniques including URL encoding the URLs on major search engines. The dynamic websites are subjected to indexing related problems and need additional SEO efforts. We also work on user behavior on the basis of users click logs.*

Keywords: *Dynamic Websites, Search Engine Optimization Technique, and Search Engine*

I. INTRODUCTION

A beautiful looking website that cannot appear in search results will fail the efforts of developer(s), because a search engine is a mean to bring a website to the searcher. A poorly optimized website would waste all efforts and money to promote it; therefore, it is a one of SEO facet to ensure a website will be indexed and ranked properly on most search engines. Current research reflects that dynamic websites face more indexing problems than static websites on most search engines. Hence, I intend to address the following questions:

What are the state-of-the-art SEO techniques for dynamic websites, and how are these techniques implemented within the three major search engines? To what extent can these state-of-the-art SEO techniques make dynamic websites search engine-friendly with regards to the three major search engines? This research work is kind of descriptive study that involves analyzing and evaluating exiting knowledge and practices in the field of web development, designing and promotion/marketing. I will mainly focus in the area of development and designing.

In this research work, I used two research methods one is literature survey and other is empirical research method to approach the two research questions addressed above. I formulated the following hypotheses to answer the experimental based of my paper [1] and [2].

Null Hypotheses:

H01*: Dynamic websites do not require search engine optimization (SEO) since three major search engines are equipped to index them as it.*

H02*: SEO techniques do not make significant difference in indexing dynamic websites on three major search engines.*

Alternative Hypotheses:

H11: *Dynamic websites require search engine optimization (SEO) for three major search engines to properly index them.*

H12*: SEO techniques make significant difference in indexing dynamic websites on three major search engines.*

Given hypotheses allowed me to make some conclusion about indexing dynamic websites on three major search engines. H01 counter H02. H02 will be evaluated only when H01 be rejected.

During experiments, I applied few SEO techniques in order to study targeted search engines' behaviors to index dynamic websites. To prove alternative hypothesis (*H11*) this can be proved that three major search engines, are not equipped to properly handle dynamic websites. So there is a need to make dynamic websites search friendly to make them visible in search results Further to determine to what extent it is possible to make a dynamic websites search engine friendly,

II. BACKGROUND TECHNIQUES

On-page SEO Techniques: On-page SEO techniques are employed on webpage(s) to optimize them to increase their worth in specific search engines. In other words, on-page SEO techniques are used to optimize factors that are related to contents of each webpage (what the users/ searchers see on webpage(s)/websites) and structure of website (what search engines crawlers see on webpage(s)/website). These techniques mainly comprise page title, header tags, Meta tags, target keywords, keywords density, ALT tags, content placement, breadcrumb trail, URL structure and size, internal linking of webpage(s), site update Frequency; last but not least, sitemaps and robot.txt files. These factors are heart of on-page SEO techniques to make website friendly for both website's users and search engines. As these techniques are important for both website's users and search engines' point of view; therefore, these techniques need to be implemented with a good care. These techniques put in the picture the theme and contents of targeted website.

Page Title/Title Tag: Page title tag is one of the most significant tags in On-page SEO because it informs both search engines and website's users about contents of particular webpage. The title tag is represented as <title> and it is basically a HTML code in the <head> section. It is important because it is used to create a string of text that appears in the top bar of Web browser. Also, search engines display page title as a headline- with hyperlink to enter your website- in search engine results. The "page title" (title tag) is essential and critical factor due to another reason because almost each search engine ranking algorithms consider title of webpage while crawling/indexing and display title in search result as well. During search engine's crawling process "page title" is a beginning point of crawlers. Moreover, searcher clicks on search result in search engine's SERP if she finds headlines/"Page Title" relevant to their search query [3] and [4].

Off-page SEO Techniques: Off-page SEO techniques are those strategies that are nothing to do with contents of website. Unlike On-page SEO techniques, these strategies are done offsite and usually these are not visible on website. Mainly, off-page SEO techniques are applied after implementation of On-page SEO techniques for improving website ranking. The process of applying off-page SEO techniques on website are not performed at once. Though, these techniques are applied gradually for better and long term result for ranking of website and for driving targeted traffic to your website. Some elements of Off-page SEO techniques have important role for improving crawl ability and Indexing of websites. In section below, I will discuss some of Off-page SEO techniques which should be considered for effective crawling and indexing of dynamic websites.

Link Building: Links building is very important element of off-page SEO techniques. During implementation of on-page SEO techniques, link structure mostly has out-bound links (which connect/link target website to other related and quality websites). These linking websites bring traffic to target website and bring positive effects on ranking of the website. However, later during implementation of off-page SEO techniques outbound and inbound both types of links are built gradually. A successful link structure may take months. There are not hard and fast rules that how many links are needed to get top ranking.The important thing that should be considered that target website is linking to quality websites (websites with high ranking). Reputed link (with better ranks) effect positively on another website and bad links can diminish website ranking. One-way link are those link which exit in target webpage but that link is not linking back to the target webpage; vice versa is called two-way linking. Both types of link are important in Off-page SEO strategies.

Search Engine for Friendly websites: The following section discusses some of the elements which become a need of website when it is needed to make easy navigation as well as easy interaction with website. On the other hand, these elements can cause obstacles in the attempt towards usability and their search engine friendliness of website. For those reasons, they need attention to optimize them if it is obliging to use elements in website design.

Optimizing Frames: Frames in website facilitate website's users to see/find contents of several webpage(s) on a single (current) webpage; this feature reduces trouble of following links; moreover, a deep hierarchy to read target page might risk that website's users will come back to the current page. In contrast, use of frame in webpage gives tough time to search engines spiders; as, frames do not correspond to the web's conceptual model [5] and [6] and [7].

Encoding URLs for Dynamic Websites URLs are entering door to webpage for search engines and website's users. Therefore, SEO experts/web developers are convinced to optimize URLs by making them short, simple and readable for both website's users and search engine point of view. There are some techniques for URL rewriting and redirecting these important to consider for both static and dynamic websites under some circumstances. The subsection below is provided with the brief details of some technique for URL encoding.

Redirecting: The use of redirects indicates both web browsers and search engine spiders that content will be fetched from new URL. Whenever, old webpage (old-page.html) be followed, it will be redirected to new webpage (new-page.html). Without redirects both website's user and search engine crawler will get 404 error (Page not Found Error). The page not found error will result the loss of traffic of website. Moreover, this

problem restricts search engines to find your webpage contents. In SEO campaign this problem can be quite paying in both cases [8] and [9] and [10] .

III. PROPOSED OPTIMIZATION TECHNIQUES

The objective and purpose of study, quality focus, perspective, better understanding and filling research gap in exciting knowledge with meaningful impact upon the field of study are some major aspects that should be included in goal definition. This experimental research is aimed to predict and determine effects of one variable on experiment by controlling several factors that may cause results to vary. My primary purpose is to use experimental research method to explore commonly used SEO techniques for improving indexing of dynamic webpage(s) on major search engines. Secondary purpose is to compare SEO techniques and see effectiveness of one technique over other. This research work leads to a better understanding; fill up a research gap in exciting knowledge and going to have meaningful impact upon the field of web development and SEO.

Experiment Planning: Experiment Planning is not an easy task; as a lot of preparation is required to plan and design experiments. A well planned experiment ensures that an experiment is carried out appropriately and that the result reflects the real world, in the best credible method.

Hypothesis: A hypothesis is an illumination for a phenomenon which can be tested in some way which idyllically whichever proves or disproves the hypothesis. Until testing of hypothesis, the hypothesis in use to be considered true and the objective of the researcher is to methodically test the term of the hypothesis. Commonly researcher use null and alternative hypothesis in the phase of formulation of hypothesis. We can describe a null hypothesis as a statistical hypothesis; null hypothesis is affirmed for acceptance. It is also called as original hypothesis.

Selection of Variables: In experimental research, variable can be described as a measurement or an attribute, whose value can vary over the course of an experiment. It helps in identifying the cause (independent variable) and effect/outcome (dependent variable) in the term of an experiment. Wrong selection of variables can effect badly on results of experiment. Therefore, it is very important to select correct variables. There are two well-known variables i.e. Independent and dependent.

In this experiment, I have selected following independent and dependent variables:

Independent Variables: SEO Techniques: Independent SEO technique or set of SEO techniques are independent variables that are used to help dynamic webpage(s) to get indexed on targeted search engines.

Independent Variables:

Dynamic Webpage(s) Indexing: Dependent variables would be the indexing of dynamic webpage(s) in targeted search engines.

Experiment Design: Experiment design is a plan to accumulate experiential knowledge, i.e. knowledge based on the analysis of investigational data. It can be practical when ensure an occurrence sequentially to improve understanding or get better performance. Making up a design means attentively choosing some experiments that are to be executing under controlled arrangements. After understanding of guidelines and unified steps, which are involved for choosing an appropriate experiment design, I chose to use The Pretest-Posttest Control Design Group as my experiment research design. This design consist two groups i.e. experimental and control. Experimental group is given the treatment and the results are accumulated at the end of experiment. The control group obtains no treatment, over the same period of experiment time, but undergoes properly the same tests.

The experiment design representation: If O1 represents pretest of experimental group, X represents the treatment and O2 represents the posttest of experimental group, then difference between O1 and O2 will be explained by X. Similarly, O3 represents pretest of control group that would not be exposed to any treatment; however, this group will be tested again after treatment. X is only applied on experimental group.

Experimental Group: Pretest →Treatment →Posttest

O1 X O2

Control Group: Pretest →Posttest

O3 X O4

In experimental group of this experiment, O1 represents static and dynamic webpage(s), which are not indexed on targeted search engine. X represents SEO techniques, which will be applied to help dynamic and static webpage(s) to get indexed on targeted search engines. O2 represents results of applied SEO techniques. In control group, O3 represents static and dynamic webpage(s), which are not indexed on targeted search engine. O4 represents results without applying any treatment (SEO technique).

Validity Evaluation: In any study, validity of results is fundamental question for researcher. All research designs are likely to have bias and design threats which can affect validity of results. Validity is the degree to which an experiment procedures what it claims to test. It is essential for experiment to be valid sequentially for the results to be precisely applied and comprehend. In the beginning researcher classify two types of validity threats. Present research expanded into in four categories.

In subsections below, I describe those validity threats, which can affect on selected research design of this experiment and how I controlled and addressed them.

External Validity: Experiment results and/or conclusion which can be validly generalized from a sample back to its parent population is said to have external validity. Due to this validity threat experiment results may not able be generalized to real world. External validity is defined in terms of Population Validity and Ecological Validity. In this experiment, it is not difficult to generalize results of this study because results are not affected if they are performed at variety of locations and there are any participants that can affect experimental data.

Population Validity: It is defined as the degree to which experiment finding and/or conclusions can be generalized from a sample back to its parent population. It is possible to generalize from study participants (subjects) or the specific finding (conclusion) based on differential characteristics. However, this validity threat is under control because no human participants are involved in the study.

Ecological Validity: It is concerns results generalization to other environmental conditions. This factor is also under control because results of my studies are not affected by variety of placed where experiments are performed.

Internal Validity: Experiment results may not be accurate because something else besides treatment X affected results of the experiments. Internal validity threats can be factors that can affect the independent variable without researchers' knowledge and affect the cause. Some factors affect internal validity includes:

History: Events which take place between observation and measurement, in addition to independent variable which affects dependent variable, are called history. It usually happens when some unplanned events occur during the experiment. During in this experimentation, I kept a complete track of optimization. I used search engine webmaster tools and server files to keep track of any unplanned events. All factors were known and such incident when server was down for one of the websites and as a result of this problem crawlers failed to indexing website.

Maturation: When subject changes between an interval T1 and T2. This often happens when people, especially children, performing a treatment on an on object change. It includes emotional, psychological, or physiological processes access time within study subject, usually individuals which operate on variables in somehow. Since it is not a study where human beings are not subject; so there was not any maturation threat.

Testing: When data gathering exercise is repeated it is likely to affect subject performance or recall. When same pretest and posttest are performed, results may affect. In my experiment, there was not repetitive pretest that could affect posttest; so this threat has no implication of validity of results. \

Instrumentation: Changes within test or data collection devise or raters may produce changes in scores on or measurements of the dependent variables. If different measures and administrator are used for pretest and posttest, results are likely to be affected. In my case, administrator did not changed throughout the process and measurement method and standard remained same throughout the process. So this threat remained under control.

Statistical Regression: If group of subjects and assembled based on extreme sores or other measurements results are likely to move towards group mean. When subject started recording at high (outliner) on a spectrum and can only go down and vice versa. There were not any outliners in this experiment so results are significant.

Selection Biases If experimental and control groups are not identical, it is likely that false variables may be affect results. In this experiment, this factor was carefully considered and experimental group was selected randomly and controlled group webpage(s) and identical in terms on data organization and basic SEO techniques.

Experimental Morality: The differential loss of study from research group is called morality. Results are affected when some subject leave in the middle of a study. It is more likely to occur in long studies. In my case, since it is a short study so there was not any human subject so this threat remained under control.

Resentful Demoralization of the control group: When control group elects to perform differently than normal, because it is not getting treatment like experimental group. In this experiment, Control group and experiment group remain alike in terms of changes and there were not any resentful changes going on control group. Only experimental group was subjected to treatments.

Fusion of Treatment: It is likely to occur in blind studies due to differences within group; it may occur due to unintentional independent variable. Since I did not perform a blind study, this threat was not significant to make any difference in results.

Statistical Conclusion Validity: It is concerned with the issue that affects the ability to draw conclusion about relationship between treatment (X) and outcome (O2). Statistical inference evaluate whether data validation procedures produce correct results so there is a need to make assure author understand statistical results performed study. In this experiment, I use SEO techniques step by step, one technique at a time by targeting particular search engines. The effect of techniques is already known when they are applied on targeted group

and results are monitored on all search engines individually. This threat is mitigated because I have prior experience of doing statistical tests so it is not very likely that I will perform statistical tests incorrectly.

Construct Validity: There are two kinds of threats to construct validity: threats related to design of experiments and threats that are concerned with social environment i.e. behavior of the participants. In this experiment, results are recorded from search engines after certain SEO techniques are applied; therefore, these results can be used to generalize the efficiency of SEO techniques for indexing of dynamic websites. Experiment design is selected after understanding of guidelines and integrated steps, which are involved for choosing an appropriate experiment design and I consider those validity threats, which can affect on research design of this experiment. Also, this experiment did not involved individuals to participate in study; so in these circumstances construct validity threat does not subsist.

Important Design Considerations for proposed Search Engine Techniques: Here we discusses of the elements which become a need of website when it is needed to make easy navigation as well as easy interaction with website.

Optimizing Frames: Frames in website facilitate website's users to see/find contents of several webpage(s) on a single (current) webpage; this feature reduces trouble of following links; moreover, a deep hierarchy to read target page might risk that website's users will come back to the current page. In contrast, use of frame in webpage gives tough time to search engines spiders; as, frames do not correspond to the web's conceptual model. Typical a webpage conceptually presents only one URL. Conversely, webpage that contains i-frames or frames aims to display contents of more than one URL (one URL for each frame) inside a single webpage. This problem poses poor impact on webpage indexing; because, search engine spider cannot follow frames contents and it might miss important contents (webpage) within frame.

If there is no other way to avoid frames then search engine recommends using alternative of framed webpage(s). The alternative of framed webpage(s) is <noframe> tag that should contain exactly same contents of <frame>; otherwise, website might be considered to spam by search engines. Search engines follow <noframe> tag and ignore actual frames. As <noframe> tag present a standalone webpage against each link/URL of frame; but it may cause navigation problem if that each webpage does not has link back to the home page. Moreover, it can cause search engines and user the no way back to the navigational page/home page. This might results that site will not be properly indexed by crawlers.

Syntax of "frames" website and alternative "no-frame" syntax of same website is provided below:

Frame Code

```
<HTML>
<Head>
<Title>Title of page goes here</Title>
</Head>
<Frameset>
<Frame SRC-"navigations.html">
<Frame SRC-"detail_page.html">
</Frameset>
</HTML>
```

No-frames Code

```
<HTML>
<Head>
<Title> Title of page goes here</Title> </Head>
<Frameset>
<Frame SRC-"navigations.html">
<Frame SRC-"detail_page.html"></Frameset>
<noframes>
Alternative of frames Code /HTML code
</noframes> </HTML>
```

Optimizing Forms: Currently, a large number of businesses rely on online buying, registration for products/services or ordering of product. Forms are popular mean of to provide interaction with websites. "Contact Us" is a simple example of webpage with form. Website's users need to fill out and submit forms. Later, on bases of filled form the resultant webpage is generated.

Unfortunately, conventional search engines' spiders cannot fill those forms. The FCF allows presentation of restricted contents in search index. It can be implemented in a way to allow searchers to see complete contents even they are registered or subscribed. However, when user clicks on original content it could

ask for registration in an honorable way. Many famous websites being use this way to index hidden form's contents by search engines. Google claim that it can fill forms but it still recommends that do not present registration forms to Googlebot instead make use of FCF for allowing crawlers to access restricted webpage.

Optimizing Flash and JavaScript: The use of Flash and JavaScript is very popular among website developers. Flash is used to create animated graphics. Web-developers most likely use them to make their website attractive. Sometimes, developers use Flash and JavaScript to create navigations which is not good for crawlers' point of view; because, Flash and JavaScript are troubling format for crawlers. Flash navigation is not spider able and results with the problem to crawl other webpage(s). Moreover, a website that heavily relays on Flash may cause website's user irritated. Search engines completely avoid JavaScript to crawl; although, <noscript> is helpful because it gives alternative contents to search engines against JavaScript.

Syntax of noscript is as follows:

```
<html>
<head> head goes here</head>
<body>
<script type= "txt/JavaScript">
document.write("This is scripting world")
</script>
<noscript> Alternative of Script goes here">
</noscript>
</body>
</html>
```

Coding URLs for Dynamic Websites:URLs are entering door to webpage for search engines and website's users. Therefore, SEO experts/web developers are convinced to optimize URLs by making them short, simple and readable for both website's users and search engine point of view.

Redirecting: Redirecting becomes necessary when contents of one website/webpage are moved to other location i.e. website/webpage. For instance, when contents of webpage
http://www.domainName.com/old-page.html has been moved to new webpage
http://www.domainName.com/new-page.html. Though, use of redirects indicates both web browsers and search engine spiders that content will be fetched from new URL. Whenever, old webpage (old-page.html) be followed, it will be redirected to new webpage (new-page.html). Without redirects both website's user and search engine crawler will get 404 error (Page not Found Error). The page not found error will result the loss of traffic of website. Moreover, this problem restricts search engines to find your webpage contents. In SEO campaign this problem can be quite paying in both cases. Therefore, redirecting is important to tell crawlers and website's users that where contents are moved for linked URL.

Redirects can be used in websites in the following situations: Hosting company is changed. CMS is changed. Old contents that are expired can be move to new. Site has broken links but has traffic and other links.

There are many ways to implement redirects. Generally they are divided into two main categories to direct crawlers, human and browsers to the new location (URL).

Methods for Redirecting: Redirects can be implemented in different ways. The subsection below is presented with some of standard redirecting method. Redirecting methods vary according to web servers. Apache web server which runs on Linux or UNIX, redirects can be implemented by using a standard files known as .htaccess and Redirect
Match directives.
The redirects that occur at page level can be implemented by Meta Tag.
Syntax of Meta tag for redirect might looks like this:

```
<Meta http-equiv"refresh" content="4"; url=http://www.doamin.com/new.html"/>
```

RedirectMatch: This method for implementation serves and facilitates same as .htaccess does. But unlike .htaccess, the Direct Match method makes use of regular expression instead of straightforward prefix matching. The provided regular expression is matched against the URL, on case of matching, the server use the new URL as file name.

Experiment Operation: The process of experimental operation can be described as the procedure of meanly defining variables into measurable elements. The procedure identifies fuzzy concepts and authorizes them to be measured, empirically and quantitatively. In this experiment, experiment operation phase includes Instrumentation, execution and data validation.

Instrumentation: In experiments, instrumentation is known as most time consuming practice. It comprises of selection of tools, documentation, and creation of guideline to perform experiments. It also contains developing pretest and posttest questions. It is very important to carefully design pretest and posttest questions; because later in the study they are subjected to answer research questions.

For instrumentation of my research experiment, it involved quite some work setting up experiment. Since I selected two websites for performing experiments, one of the website had to be a new dynamic website so that I can closely monitor, how search engines respond to a new website in terms of indexing its webpage(s). I had to create a sample dynamic website for this purpose.

I also planned to examine effectiveness of URL rewriting on search engines. For this purpose I planned to use a websites where I can easily tryout URL rewriting. It is effective to use URL rewriting in early stages of website development because it is quite some work to implement URL rewriting on a developed website. Another reason to keep URL rewriting on a different website was to control any possible influence on newly uploaded website. So I used a website that was developed in PHP template engines which allows creating friendly URL with less effort. There are three types of instruments, which are known as common for using in experimentation i.e. Objects, Guidelines and Measurement tools. Following of detail description of instruments which I used in this experiment:

Objects: In computer science experimental research, object can defines as specifications or code documents which are used to execute experiments. Objects used in this experiment are webpage(s) that are subjected to indexing on targeted search engines. Also search engine's SEO techniques that are specific to dynamic websites and common for static and dynamic websites.

Guidelines: Before execution of experiment it is necessary for researchers that they are well aware about guidelines to execute the experiment. Also, they must need to know about topic and its background, aims and scope of experiment in depth. For this experiment, I reviewed several research papers to familiarize myself with experimental research and writing technical documentation. I also performed research on how to perform experiments for investigative studies. After understanding of experimental research, I decided on scope of chosen topic and prepared guideline for myself.

Measurement Tools: The tool used for measurements and data collection on a variable is called measurement tool. In this experiment, I used webmaster tools for Google, Yahoo, MSN, and Bing to perform SEO techniques. I also used Google Analytics account to keep record on traffic and monitor traffic reports. Other than these online tools, I also recorded search results from SERPs of each Search on daily basis and used excel sheets to record collected data.

Execution: After an instrumentation process has been done, the next step is execution of experiment. This experiment is conducted in a period of two month after development and optimization (SEO) of both websites which, I am using in this experimental research. Initially, it was not known how long search engines will take to index webpage(s) (experimental group) because it depends upon search engine's results. However, I intended to complete experiment in a period of 30 days.

Since experiment required online websites so there was no need to have fix setting or place to take experiment readings. To avoid SEO techniques results overlapping, I applied and evaluated SEO techniques step by step. I used two websites in this experiment; one websites was new to the search engines and seconds has web presence since last one year. Main purpose to use two websites of different ages was to study, if search engines response to them differently or not. First, I created a new dynamic website and uploaded it on server so that I can study behavior of search engines and study if search engines can index the target website without setbacks. Second, I applied SEO techniques, one technique at a time, on webpage(s); I was checking search engines' results every 6 hours. However, I was recording experiment results only once in a day. There was not any need to records results frequently because results were not very likely to vary but I was taking my reading according to my plan. There was only one instance that results were seen changed within six hours after performing one technique in Bing webmaster tool. Results were recorded in Microsoft Excel sheets and manual search results were recorded as screenshots taken from webmaster tools and search engines

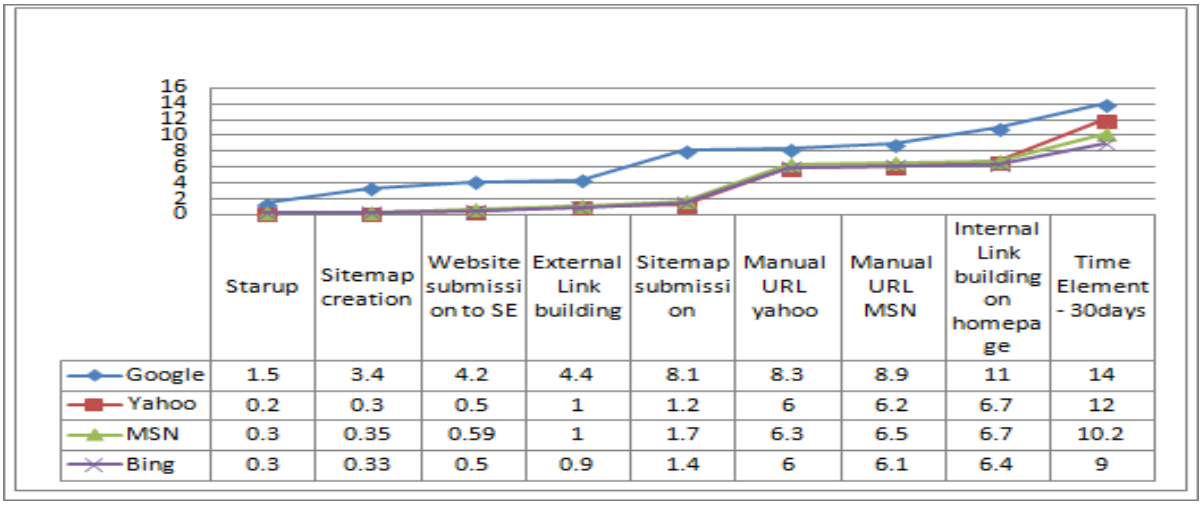
. EXPERIMENTAL RESULTS

Results were recorded in two steps. In first steps, I applied SEO techniques step by step and experiment results were recorded on daily basis. For this step results were last taken after applying last SEO technique. Then there was wait period of about thirty days to study effect of time element on search engine's indexing mechanism. I did not apply any additional technique during this period; however, I was keeping track of server log files to make sure there are no underlying server side problems that may affect experimental data. In second step, final results were taken for targeted search engines. Although results were recorded on daily basis, results presented here are summarized to demonstrate effectiveness of each SEO technique useful on targeted search engines. In the following sub-sections, I will present results for both websites.

Website1: wasfabththesis.com: Website1, a relatively small website, contains total 21 URLs (including 18 dynamic and 3 static webpage(s)). I aimed to index these webpage(s) on search engines by applying certain SEO techniques sequentially; website's pages other then targeted 21 webpage(s) were access denied by crawlers. At the end of experiment total 72% webpage(s) were indexed in Google, 44% webpage(s) were indexed on Bing

and35% MSN and Yahoo, and 30% webpage(s) were indexed on Ask and AOL Search engine including all static webpage(s).

Figure 1 is descriptive representation of indexing throughout experimentation with step by step application of SEO techniques for website1 (wasfabththesis.com).



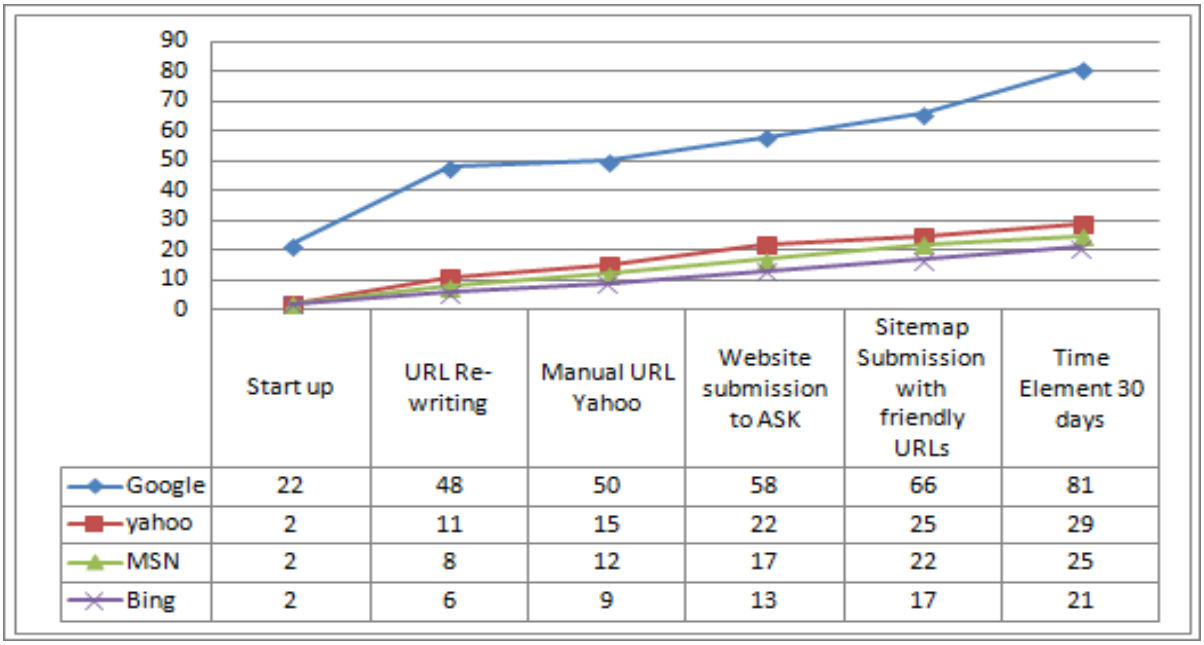
	Starup	Sitemap creation	Website submission to SE	External Link building	Sitemap submission	Manual URL yahoo	Manual URL MSN	Internal Link building on homepage	Time Element - 30days
Google	1.5	3.4	4.2	4.4	8.1	8.3	8.9	11	14
Yahoo	0.2	0.3	0.5	1	1.2	6	6.2	6.7	12
MSN	0.3	0.35	0.59	1	1.7	6.3	6.5	6.7	10.2
Bing	0.3	0.33	0.5	0.9	1.4	6	6.1	6.4	9

Figure1 Descriptive representation of indexing throughout experimentation with step by step application of SEO techniques for website1

This website was uploaded on server on the day when I started experiment so that I can watch indexing process. Before applying a SEO technique, a sample of webpage(s) was randomly chosen from the webpage(s) that were not indexed. For simplicity, indexing results are summarized. Each entry represents number of webpage(s) indexed after applying step by step SEO technique on experimental group. Efficiency of each SEO technique is also calculated individually to determine how effective each technique.

Website2: recipe-planner.com: This website contains 131 dynamic webpage(s) with 38 webpage(s) containing friendly URLs which served as our experimental group to test effect of friendly URLs on search engines. This website site had web existence last one year with 18% webpage(s) indexed in Google. At the end of this study percentage increased to 57% on Google. Furthermore, a total of 20% webpage(s) were index in Yahoo, 18% in Bing, 15% in AOL and 13% were recorded in Ask.

Website2 contained 125 dynamic pages. This website had web presence since last one year but 23 webpage) of website were indexed in Google. Home page of website was indexed in Yahoo, AOL and Bing. However, website was not index in Ask. After applying several SEO techniques website varied in all search engines. Following graph is descriptive representation of indexing throughout experimentation with step by step application of SEO techniques for website2.



	Start up	URL Re-writing	Manual URL Yahoo	Website submission to ASK	Sitemap Submission with friendly URLs	Time Element 30 days
Google	22	48	50	58	66	81
yahoo	2	11	15	22	25	29
MSN	2	8	12	17	22	25
Bing	2	6	9	13	17	21

Figure 2 Descriptive representation of indexing throughout experimentation with step by step application of SEO techniques for website2

Friendly URLs: As discussed that some experiment on most controversial variable i.e. user friendly URLs; because I aimed to study how effective this SEO technique is for indexing URLs and how the search engines respond to this technique. This is second part of experiment where I examined if friendly URLs help improve indexing on search engines because unnecessary parameters are not included in the URL. As discussed in literature review, Google claims it does not has any problem reading dynamic looking URLs and it is not required to rewrite URLs to eliminate parameters. Two samples were taken from the population and URL rewriting (treatment) was applied on experimental group and results were recorded in Excel sheet and basic analysis was performed automatically.

IV. RESULTS ANALYSIS

I created quantitative assessment of effectiveness of each SEO techniques applied on both websites used in research. To achieve this objective, I organized and summarized data collected during experimentation so that I can analyze and evaluate what I have discovered and present in a form that a conclusion can be derived from it. First, I summarized data and recorded search results taken from search engines and webmaster tools which I used in experiment. This process cost me quite some time because I recorded all data manually as there were not any tools available that I could use to get summary index reports from all search engines. Second, I summarized data in simplest form so that results can be evaluated and described for corresponding SEO technique. Finally, I summarized data so that I test hypotheses and suggest for significant findings.

V. CONCLUSION AND FUTURE WORKS

SEO is essential for websites to increase their visibility in search engines. On-page SEO techniques and Off-page SEO techniques are commonly used SEO techniques for websites. On-page SEO techniques help in optimizing website contents by making it accessible/understandable for search engines; because search engines can understand only certain formats of website contents. Whereas, Off-page SEO is nothing to do with website contents but it affects websites to get better ranking and improve crawlers' visit ratio to the websites. The experimental results showed that URL encoding of dynamic webpage(s) improved indexing results by 34.12% and 31.87% for Yahoo and Bing respectively. These are significant figures to reflect the effectiveness of URL encoding URL/friendly URLs on these search engines. Conversely, Google result does not reflect such positive (though not negative) response towards friendly URLs. The URL encoding is highly recommend not only because URL encoded webpage(s) are more likely to indexed in Bing and Yahoo but also because friendly URLs are more likely to appear in top results in Google. Since a properly indexed websites are an initial step towards better ranking to bring websites more visible in search engines, it is important to optimize a websites properly to increase credibility of a websites. In future, I would like to perform a detailed study on techniques that can help achieve a real website acquire higher ranking. I would also like to explore that whether it is possible to achieve top/better ranking in Google, Yahoo and Bing at the same time by using some common SEO techniques.

REFERECES

[1]. G. Rogan; "A study of Search Engine Optimization methods", National University of Ireland, Gateway (2009).

[2]. N. Nazar;, "Exploring SEO Techniques for Web 2.0 Websites", Master of science Thesis in software Engineering and Technology, Department of CS and Engineering Chalmers University Of Technology Goteborg, Sweden, (June 2009)

[3]. R. Vadivel & Dr. K. Baskaran; "Implementing Search Engine Optimization Technique to Dynamic/Model View Controller Web Application", Global Journal of the Computer Science and Technology Vol. 10 Issue 6, V1.0, (July 2010)

[4]. C Duda, G Frey, D. Kossmann & C. Zhau;, "AJAX Search: Crawling, Indexing & Searching Web 2.0 Applications", Prfocedding of the VLDB Endowment v.1 n.2, (2008)

[5]. X. Zhang, M. Zuo, & Q. Liu;, "Analysis of the Reasons Why Invisible Web Can't Be Seen and its Effective Retrieval Strategies", Innovative Computing Information and Control (ICICIC) '08 3rd International Conference on, vol., no., pp.563-563, 18-20, (June 2008)

[6]. M. Alfano & B. Lenzitti;, "A Web Search Methodology for Different User Typologies" International Conference on Computer Systems and Technologies-CompSysTech'09, (2009

[7]. C. Sheram; "Google Power: Unleash the full potential of Google", McGraws-Hill /Osborne Companies, ISBN 0-07-225787-3, (2005)

[8]. Y. Ru & E. Horowitz; "Indexing the invisible web: a survey", Department of Computer Science, University of Southern California, Los Angeles, California, USA page 2, (2005)

[9]. L. Huang;"Challenging the Invisible Web: Improving Web Meta-Search by Combining Constraint-based Query Translation and Adaptive User Interface Construction", The Department of Computer Science of Technische University Darmstadt, Germany, (2003)

[10]. J. Köhne "Optimizing a large dynamically generated website for search engine crawling and ranking", Department of Media- and Knowledge Engineering Technical University of Delft Netherlands, (2006).

Author1-Mr. Mahendra Pratap Singh Dohare has obtained M.C.A. degree from MITS Gwalior (Under RGPV University, Bhopal MP) in 2009 and after obtained M.Tech. (CSE) degree in Computer Science with Specialization in Software System from SATI, Vidisha MP in the year 2011. He has Qualified GATE Exam Thee Times in the year (2008, 2010, & 2012) in the Subject of Computer Science & Engineering and also JAM Exam Qualified In 2005 in the Subject of Mathematics. Presently he is pursuing Ph.D. in Computer Application from MITS Gwalior (Under RGPV University, Bhopal MP). His research interest includes Data mining, Computer Network, Operating System & Software Testing. He has authored and co-authored 06 research papers in national, international, journals, Conferences, Seminars and Proceeding. Currently, he is working on to develop web Optimization Using Web Mining Techniques with User Profiling in on line searching for Click Stream data processing in data mining.