



## **Thesis Report**

### **“Efficacy Measurement and Analysis of Google Scholar, as a Search Engine for Scholarly Publication”**

B.Sc. in Electronic and Telecommunication Engineering  
Department of ECE  
East West University, Dhaka  
Bangladesh

#### **Submitted By**

**Munim Mowsum**

ID: 2013-1-55-028

**Mohyminul Islam**

ID: 2013-2-55-002

**Md Sohan Prodhan**

ID: 2013-2-55-003

**Under the supervision of**

**Dr. Mohammad Arifuzzaman**

**Assistant Professor**

**Department of Electronics and Communications Engineering**

**East West University**

# DECLARATION

We hereby declare that we carried out the work reported in this thesis in the Department of Electronics and Communications Engineering, East West University, under the supervision of Dr. Mohammad Arifuzzaman. We solemnly declare that to the best of our knowledge, no part of this report has been submitted elsewhere for award of any degree. All sources of knowledge used in this report have been duly acknowledged.

.....

Munim Mowsum

ID: 2013-1-55-028

.....

Mohyminul Islam

ID: 2013-2-55-002

.....

Md Sohan Prodhan

ID: 2013-2-55-003



# CERTIFICATE

This is to certify that the thesis entitled “Efficacy Measurement and Analysis of Google Scholar, as a Search Engine for Scholarly Publication” , being submitted by Munim Mowsum, Mohyminul Islam and Md Sohan Prodhan Department of Electronics and Communications Engineering, East West University, Dhaka in partial fulfillment for the award of the degree of Bachelor of Science in Electronics and Telecommunication Engineering, is a record of major thesis carried out by them. They have worked under my supervision and guidance and have fulfilled the requirements which, to my knowledge, have reached the requisite standard for submission of this dissertation.

.....

Dr. Mohammad Arifuzzaman

Assistant Professor

Dept. of Electronics and Communications Engineering



# Approval

This is to certify that the thesis entitled “Efficacy Measurement and Analysis of Google Scholar, as a Search Engine for Scholarly Publication” submitted to the respected member of the faculty of Engineering for partial fulfillment of requirement for the degree of Bachelor of Electronics and Telecommunication Engineering by the following student and has been accepted as satisfactory.

Submitted By :

**Munim Mowsum**  
ID # 2013-1-55-028

**Mohyminul Islam**  
ID # 2013-2-55-002

**Md Sohan Prodhan**  
ID # 2013-2-55-003

.....  
**Dr. Mohammad Arifuzzaman**  
Assistant Professor  
Dept. of Electronics and Communications Engineering  
East West University

.....  
**Dr. M. Mofazzal Hossain**  
Chairperson & Associated Professor  
Dept. of Electronics and Communication Engineering  
East West University



# Acknowledgement

We would like to express our gratitude to our supervisor Dr. Mohammad Arifuzzaman, Assistant Professor, Dept. of Electronics and Communications Engineering, East West University for his guidance and support throughout this thesis work. He has been a constant source of inspiration to us throughout the period of this work.

Our appreciation goes to our parents who supported us all these years. Their unconditional love, encouragement and inspiration gave us the strength to complete this thesis.

We are grateful to Chairman, all faculty members of the Department of Electronics and Communications Engineering as well as the concerned officials for their cooperation.

We would also like to thank Alok Paul, graduate teaching assistant of East West University for his support in completing this thesis.

.....  
Munim Mowsum  
ID: 2013-1-55-000

.....  
Mohyminul Islam  
ID: 2013-2-55-002

.....  
Md Sohan Prodhan  
ID: 2013-2-55-003



# Abstract

This article introduces the concept of functioning Google scholar as a search engine for scholarly publication. We have done a brief analysis on Google scholar ranking algorithm, which explains the mechanism of Google scholar search engine in finding the expected search in particular academic arena. In Google scholar the scholarly documents are indexed and user who eager to search for a literary work, always try to retrieve their document of interest using Google scholar search engine. Actually, the basic functionalities of general purpose search engines and academic search engines is different. And that is the main reason where Academic Search Engine Optimization (ASEO) and its working algorithm should be more precise. Though we mainly focus on Google Scholar Search Engine, we also cover the basic properties which are strongly considered behind the algorithm of Academic Search Engine. In addition to that, we also include some potential recommendations for academic search engines, which will contribute to develop more effective academic search engine algorithm in coming days.

# INDEX

<b>Contents</b>	<b>Page No</b>
Declaration .....	i
Approval.....	ii
Certificate .....	iii
Acknowledgement.....	iv
Abstract.....	v
 <b>CHAPTER 1:</b>	
Introduction.....	2
 <b>CHAPTER 2:</b>	
An Overview of Academic Search Engine.....	4
 <b>CHAPTER 3:</b>	
Analysis on Google Scholar Ranking Algorithm.....	8
<i>3.1 Citation Count</i> .....	9
<i>3.2 Date of Publication</i> .....	12
<i>3.3 Number of Patents</i> .....	16

## **CHAPTER 4:**

Findings.....	23
---------------	----

## **CHAPTER 5:**

Conclusion and scope of future work.....	26
--	----

<b>REFERENCES.....</b>	<b>28</b>
------------------------	-----------

## **LIST OF FIGURES**

Figure 2.1 : Google Scholar (Homepage).....	5
Figure 3.1 : Citation count per position and page.....	10
Figure 3.2 : Date of Publication per position and page.....	15
Figure 3.3 : No. of Patents over 10 pages .....	18
Figure 3.4 : Search Page 1 (keyword : search engine optimization).....	20

## **LIST OF TABLES**

Table 3.1 : Citation count table.....	11
Table 3.2 : Date of Publication table.....	14
Table 3.3 : No. of Patents table.....	19



# CHAPTER 1



## Introduction:

Researchers have an interest of publishing their articles in various renowned journal, magazine and proceedings. Researchers have contemporary article and innovation. To continue their high quality research a perfect search engine is necessary [1]. That's why academic search engines are needed such as Google scholar [2], Microsoft Academic search [3] etc. But there are thousands of papers keeping out of order in the web. Academic search engines take the job to gather the scholarly documents for researchers. Then the big challenges become for a researcher to know about the optimizing function of that particular academic search engine. Because the reader always try to find out their expected result on the first page. In a word, articles positioning in the top are more preferred. This article is an initiative to find the key concept about the function of Google scholar search engine. This article correlate the page ranking of Google scholar with several parameters like citation count, date of publication, author's reputation, journal's ranking etc. In this article we have discussed about overview of academic search engine which plays a vital role for every web pages. This is briefly discussed about how citation count effects to be top ranked in the search engine the more citation count, the value of paper will increase along with author's importance. The rest of the paper is organized as follows: In chapter 2 an overview of academic search engine is discussed. In chapter 3, Google scholar ranking algorithm is analyzed.

# CHAPTER 2

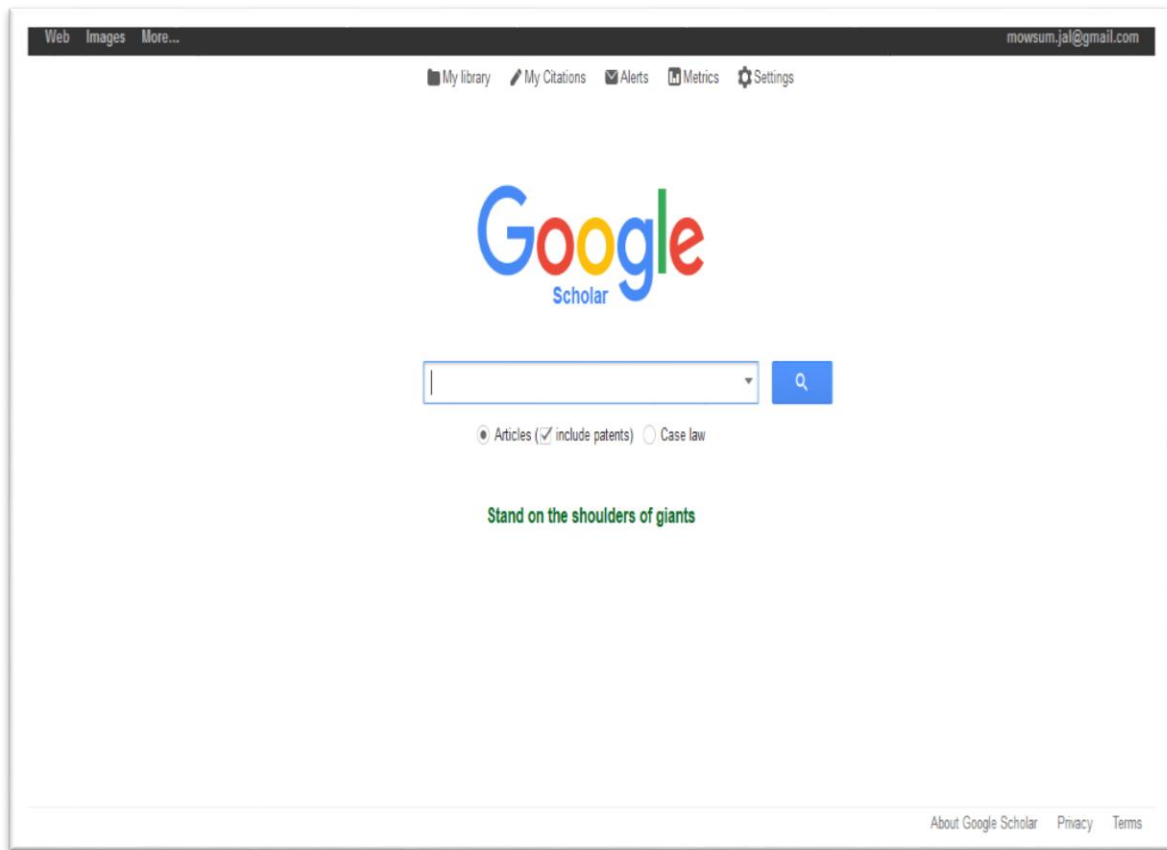


## 2.1 An Overview of Academic Search Engine:

The fundamental concept of Academic search Engine is the same for all. There are different ranking algorithm work in the backend operation of the search engine to display in which position the result will be published .There are some common keywords like citation count, date, author's value etc. work behind the operation of Academic search engine. [4]There are several types of academic search engines such as Google scholar, Microsoft Academic search , Info mine, Internet Public Library, Academic Index, and Digital Library of the Commons Repository etc.

[1] Academic Search Engine Optimization also known as ASEO does several important jobs. It helps to show the related or preferable scholarly documents at first. To do this, any implementation and modification are needed then that is also done by Academic Search Engine Optimization (ASEO).

Academic Search Engine is only for the ones who try to search especially scholarly documents. Additionally, for people searching of scholarly documents in Google scholar, academic documents are organized, labeled and formatted completely for them.



**Figure 2.1: Google Scholar (Homepage)**

## **2.2 Academic SEO overview:**

Four fundamental building blocks of academic SEO ,how to optimize individual document types (PDF, WORD, and HTML etc.) The role of metadata is creating and optimizing way finders for academic literature. There are some Common myths and misconceptions about meta data.

[5] Metadata is talking about the data which is focused on the other data. Metadata internally drives the search engine optimization. When any topics is given in the search bar for searching in particular academic search engine then it tries to communicate important

information inside the web pages directly .In a word, metadata works behind the scenes. Academic search engines use different algorithm to display original searched item.

There are some common misconceptions about SEO. [6]

1. SEO is all about keywords and links .Keywords and links certainly play a role in SEO, but they are not the only factors.

2. Keywords (not provided)' means the end of SEO .Instead of focusing on the keywords, focusing on the internal content will be more fruitful.

3. Sub-headers are important for on-page SEO.

# CHAPTER 3



### **3. Analysis on Google Scholar Ranking Algorithm:**

Google Scholar considers many factors behind its search results. These factors are quite similar to the usual algorithm of Google search engine or other search engines. Together these factors keep impact on search results and determine the effectiveness of the outcome.

Some of the factors considered in Google's algorithm are citation count, articles age, search term occurrence in an article's full text, search term frequency in an article's full text, search term occurrence in an article's title, search term occurrence in author or publication name and so on [8]. Three of the key factors we will discuss are citation count, date of publication and Number of patents. These three factors altogether have tremendous impact on finding suitable articles from Google scholar or other academic search engines.

Google Scholar gives some option explicitly during searching. It shows different timeline and gives option whether it will include article which has patents and citations. These filter options make searching a lot more easily for the user. Still these are not the only options that are related to the whole process. Some of the results can be acquired by taking total measures using these options.

In the later part of this article some evidence related to patents will be shown.

Values and graphs based on citation count, date of publication and patent for the search keywords 'search engine optimization' has been illustrated and some analysis based on the result has been done.



We will also give a deep look on some other factors which we consider very appropriate in terms with the Google scholar's efficacy.

### 3.1 Citation Count

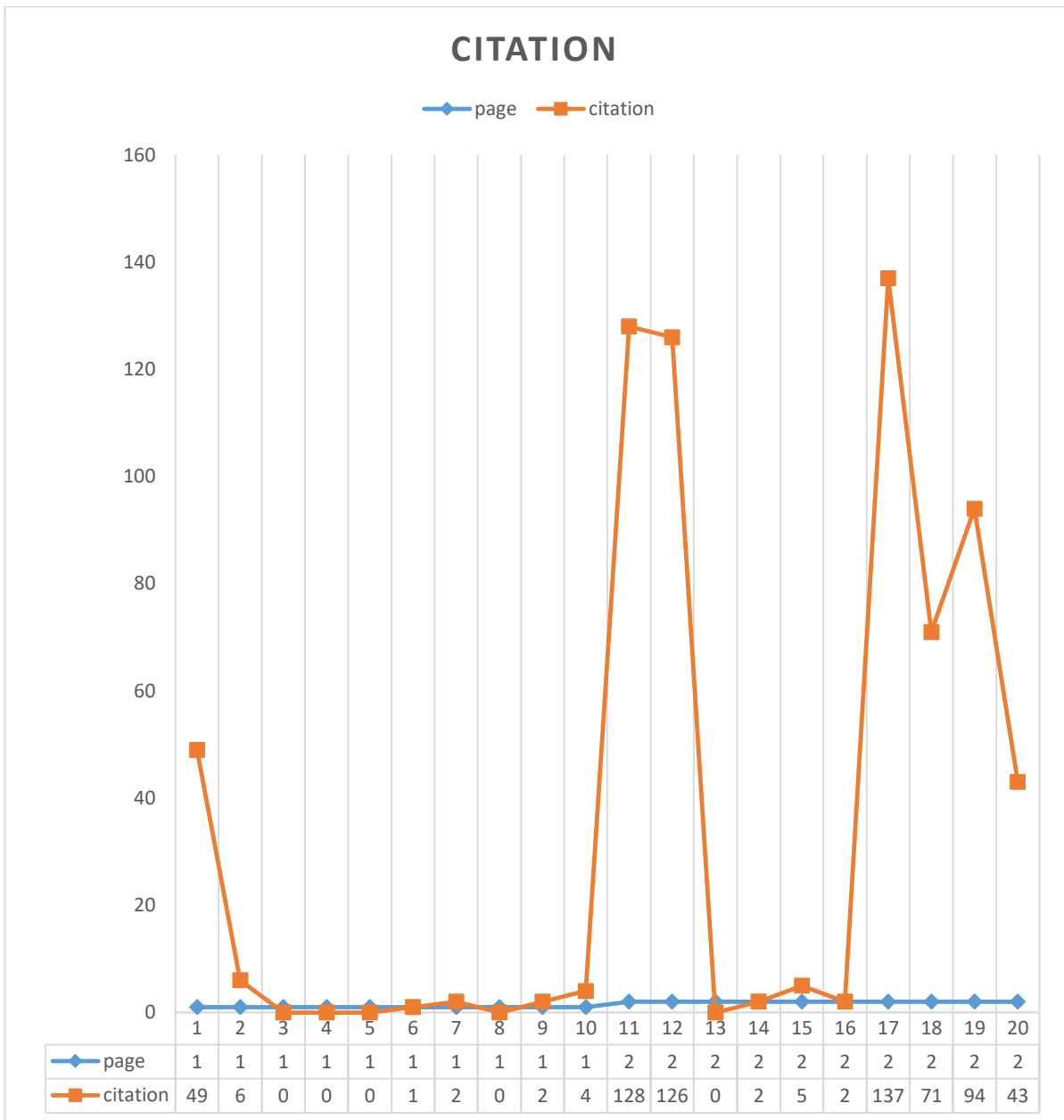
Giving citations major and common practice in journals. Citation is the way of telling readers that certain material in the work came from another source .It also helps to dig more based on other references.

An article which has more citation simply leads it as a more useful article than others. Based on the citation author's h-index<sup>1</sup> is defined. However in many cases article with highest citation won't be the topper of the search results [9]. Because some of the other factors may come across and due to their bad impact an article with good citation goes down to the list. Such as if some particular articles have higher citation than other related articles but the search keys frequency is low in these documents, so they may stay behind the line. Hence, citation count isn't the ultimate parameter rather a parameter with huge significance.

For the word “ Search Engine Optimization” 20 search results have been taken.Then we've taken their citation count . By the graphical illustration the highs and lows of the graph shows the dissimilarities and unconventional pattern of citations count.

---

<sup>1</sup> The **h-index** is an author-level metric that attempts to measure both the productivity and citation impact of the publications of a scientist or scholar. The index is based on the set of the scientist's most cited papers and the number of citations that they have received in other publications. The index can also be applied to the productivity and impact of a scholarly journal[10] as well as a group of scientist.



**Figure 3.1 : Citation count per position and page**

Position	Page	Citation
1	1	49
2	1	6
3	1	0
4	1	0
5	1	0
6	1	1
7	1	2
8	1	0
9	1	2
10	1	4
11	2	128
12	2	126
13	2	0
14	2	2
15	2	5
16	2	2
17	2	137
18	2	71
19	2	94
20	2	43

**Table 3.1: Citation count table**

In figure 3.1, page no. is shown in blue label and citation count is shown in red label. Citation count is shown in terms of position and page. From table 3.1 , it can be seen that , in search position 1 , the number of citation count is 49 , it decreases much and citation count becomes 6 at position 2, then there's consecutive three 0 citation followed by a

single citation at position 6. Then it is noticeable that the ups and downs of citation count continues. At 11th and 12th position there's a sudden rise of 128 and 126 citation counts respectively and then it goes down to 0. At No. 17<sup>th</sup> position the top citation count of 17 is recorded.

From figure 3.1 and table 3.1, it can be seen that the pattern of citation count is quite unlikely comparing with the conventional search engine results. We get the highest citation count at second page and at 17<sup>th</sup> position. But citation count is an important parameter and it is being neglected. User expects to find articles with higher citation in the first page of search results. But here the density of citation count is higher in the second page than the first page. That surely a bizarre search results and cannot be expected by a user in any circumstances.

### **3.2 Date of Publication**

Date of publication is a major consideration for a search result. Outdated articles aren't to be expected in top position. Articles which are latest and have healthy citation are expected at top position. Also there are multiple editions of some articles or articles that include lot more new research topics than their previous copy are expected at front.

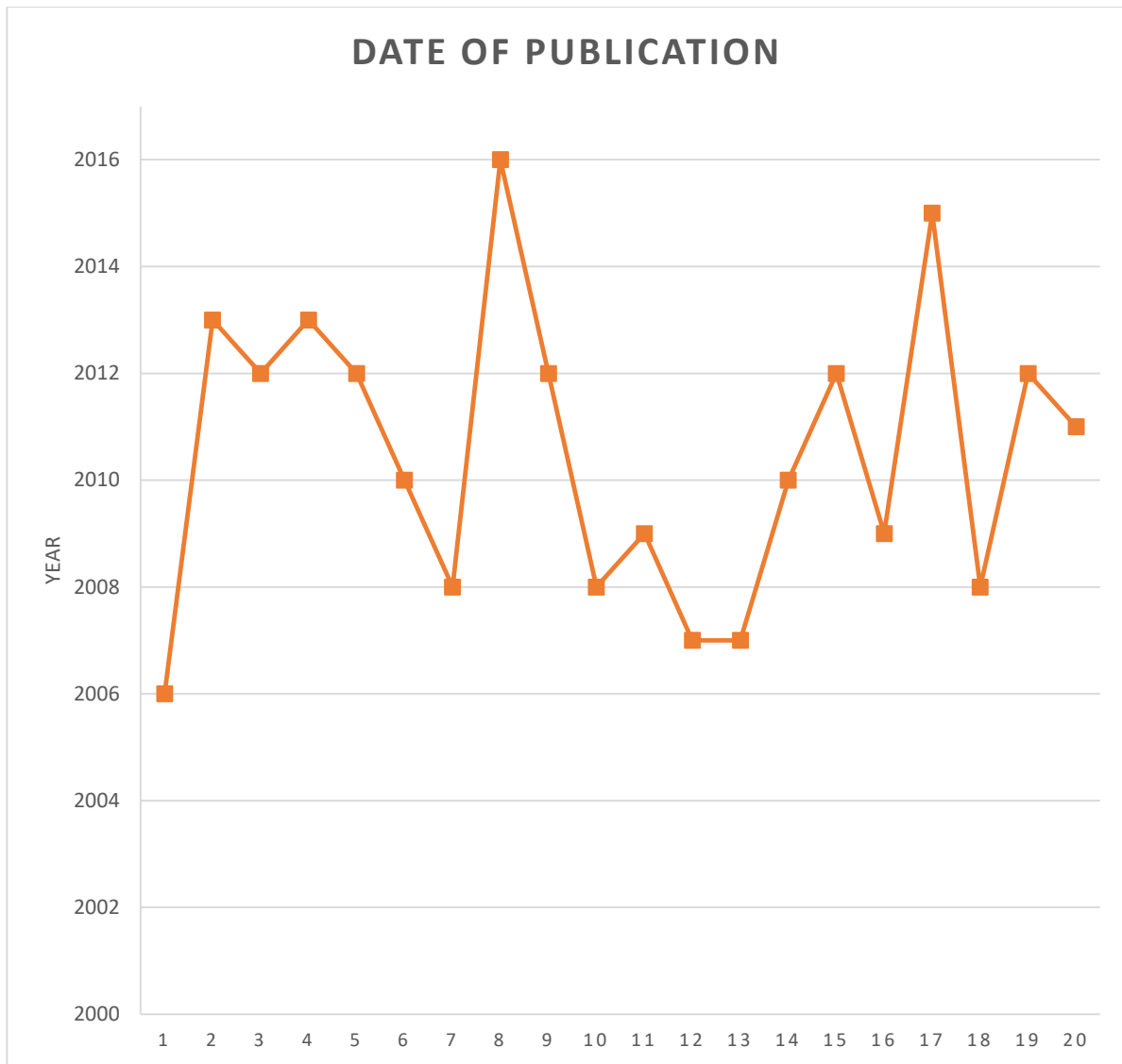
If an article is published a few years back in contrast with others which have much higher citation count than others then it is quite likely to be shown at top position. An article with patent or an article which is revolutionary in it's field is also to expected at the highest position. Articles which are written by famous researchers from very reputed university also expected to be a top level entry.

It has to be noticed from table 3.1 and figure 3.2 that the date entry is started from 2000. Google hasn't included many articles prior to the 2000. This is quite understandable, because maximum older articles aren't available in online. A significant fact has to be mentioned that articles published in a same year or two consecutive years shouldn't be allowed in a large amount in a single page. Because that isn't conventional.

Figure 3.2 and table 3.2 have shown that the articles publishing year is between 2006 and 2016. The span is 10 years. It has included two articles at year 2012 and two articles at year 2013. Therefore, a considerably good amount of articles have been got in 10 years span quite thoroughly. At the topmost position we get the article published in 2006 and then followed by an article which is published in 2013. Comparing their citation the first article got 49 citations and the second article got 6 citations. So, the results are quite expected. Then if we take a look at next three consecutive articles published in 2012, 2013, 2012 respectively and they have zero citations. At 6<sup>th</sup> and 7<sup>th</sup> position the publication dates are 2010 and 2008 respectively. The 2010's one got a single citation and the other got 2 citation. As they are older articles and have a lesser citation count, they are placed at lower position in the search results.

At the second page, in 11<sup>th</sup> and 12<sup>th</sup> position we get 128 and 126 citations respectively and their publishing years are 2009 and 2007 respectively. Though they are a bit older, they've got a huge amount of citation count. It's about 2.6 times higher than our top most article's citation count. Based on this evidence it can at least be expected that this two article deserve their position in the first page of the search results. The Highest citation count is found at 17<sup>th</sup> position. From the Date of publication table, it's publication year recorded as year 2015. Therefore, it is quite latest article with the maximum number of citation count. It has to be placed at the top of our table; rather it comes at a much lower position. This result is quite surprising and not to be expected at all. If we compare all of these figures, tables we can see that there's a huge unusual pattern at 11<sup>th</sup>, 12<sup>th</sup> and 17<sup>th</sup> positions.

These articles are quite better than the other top search results, still some articles with least citation count and also in terms with publishing factor stays in good positions.



**Figure 3.2 : Date of Publication per position and page**

Position	Page No	Date
1	1	2006
2	1	2013
3	1	2012
4	1	2013
5	1	2012
6	1	2010
7	1	2008
8	1	2016
9	1	2012
10	1	2008
11	2	2009
12	2	2007
13	2	2007
14	2	2010
15	2	2012
16	2	2009
17	2	2015
18	2	2008
19	2	2012
20	2	2011

**Table 3.2: Date of Publication table**

The unusual pattern of our results raise a question mark whether the algorithm of Google scholar's search engine is as effective as it meant to be. It under whelms the two key factors: Citation Count and Date of Publication.

As the graphical representations of both citation count and Date of Publication indicates that there's no pattern followed in the search results based on this factors. In later part the No. of patents amongst these articles have been shown. It will be further discussed that, based on the three major conditions of an academic search engine, how effectively Google scholar works, what problems it is facing based on our search results.

### **3.3 Number of Patents**

A patent is an exclusive right granted by a country to an inventor, allowing the inventor to exclude others from making, using or selling his or her invention in that country during the life of the patent. It does not give the inventor the right to use or "practice" the invention, and thus the right is subject to any prior rights that others may have to related inventions [11].

Patent is a very touching and important issue for a researcher. Every outstanding invention demands a patent of it's own. It also gives recognition to the inventor and brings huge prestige for the individual's work.

Google has it's own patent search engine-Google Patents. Google Patents ([www.Google.com/patents](http://www.Google.com/patents)) includes over 8 million full-text patents. Google Patents works in the same way as the Google search engine.

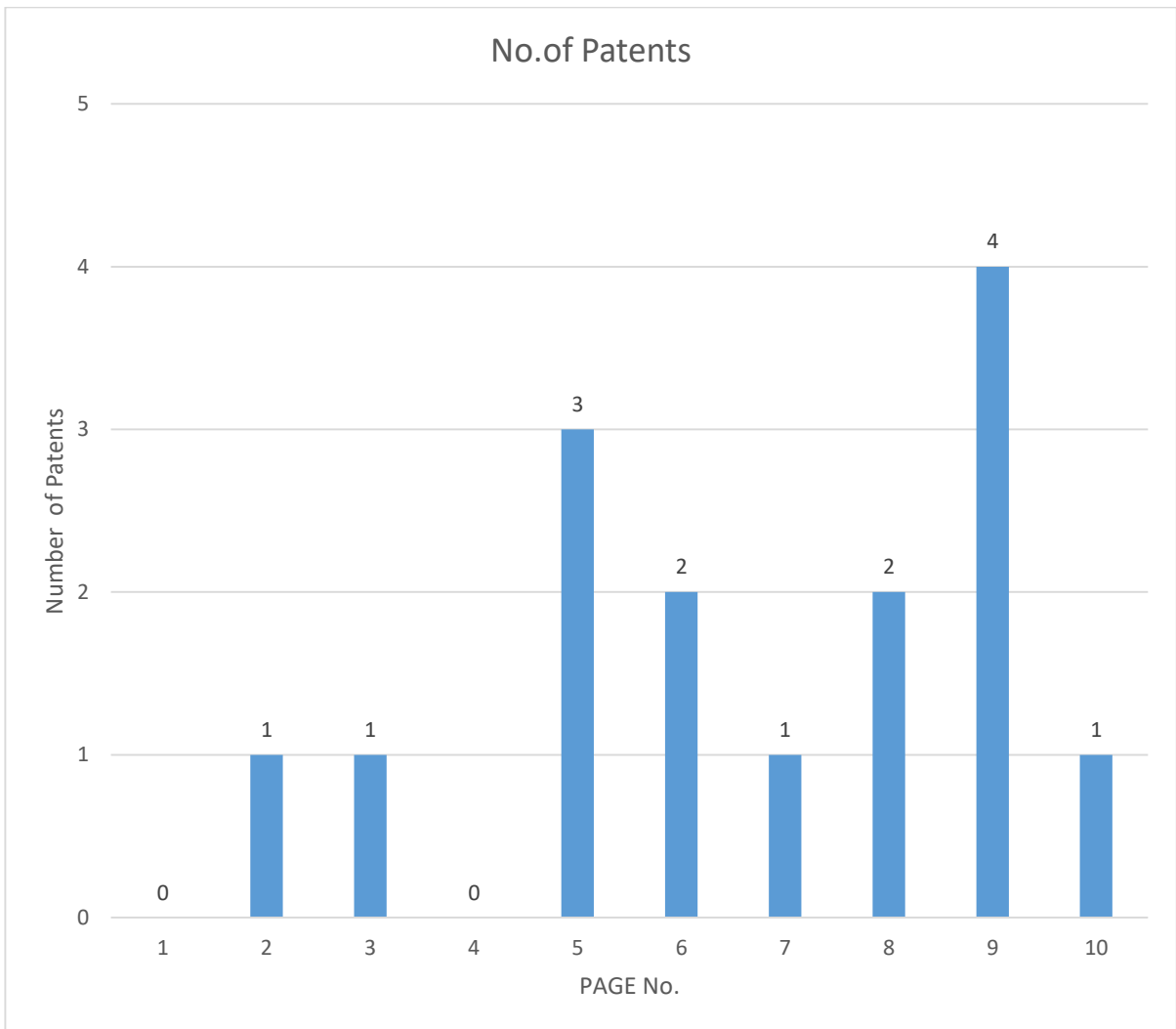


Google Patents is the global patent search engine that lets users search through patents from the USPTO (United States Patent and Trademark Office), EPO (European Patent Office), etc.[12].But in our study we won't consider Google Patents. Because it is an academic search engine based on a particular parameter. Not many parameters are considered in Google Patents. However, as a fully-functioned academic search engine Google Scholar needs to consider patent as an important issue.

Therefore, it is crucial for an academic search engine like Google scholar to include the articles consists of patent into it's topmost search items. Here, for the word 'search engine optimization' amount of patents over 10 pages have been taken and some unexpected results have been found. The lowest number has been found at the first page and it is null. The highest number has been found at the page 9 and it is 4.

Second and Third page of our search result consists of a single patent each. Page 4 doesn't contain any article which has any patents. Page 5, 6 and 7 have 3, 2& 1 numbers of patents respectively. While Page 8 consists of 2 articles having patents and Page 10 got a single article having a patent. The concentration of patents in the relative pages is quiet strange. Page 1 to 5 got five articles which have patents, on the contrary page 6 to 10 there are ten articles which have patents. That is exactly double in contrast with the first five pages. Actually it should've appeared the vice-versa.

The first page and the fourth page doesn't include any article consists of patents. The fifth and the ninth page consist of three and four article respectively which got patents.



**Figure 3.3: No. of Patents over 10 pages**

The result doesn't satisfy the user's demand of getting higher concentration of articles which have patents.

Page No.	Number of Patents
1	0
2	1
3	1
4	0
5	3
6	2
7	1
8	2
9	4
10	1

**Table 3.3: No. of Patents Table**

Looking at the all table and figures ,it has been found that most citation count has been existed at the second page; articles which have relatively new publication date in contrast with citation also stays at the second page and now it has also been seen that highest

concentration of patents also found in the later pages. None of these findings resembles with the expected findings. Rather, it seems to be obvious Google scholar doesn't give importance of our considered parameter in it's algorithm.

The screenshot shows a Google Scholar search interface. At the top, it displays 'Scholar' and 'About 1,730,000 results (0.03 sec)'. On the right, there is a 'My Citations' button. The main content area lists search results for the keyword 'search engine optimization'. The first result is a book by H. Davis from 2006, with a snippet: 'SEO--short for Search Engine Optimization--is the art, craft, and science of driving web traffic to web sites. Web traffic is food, drink, and oxygen--in short, life itself--to any web-based business. Whether your web site depends on broad, general traffic, or high-quality, targeted'. It is cited by 52. The second result is a thesis by T. Jež from 2012, with a snippet: 'Abstract This diploma thesis deals with the analysis of optimizing websites for search engines (Search Engine Optimization--SEO), summarizes history and evolution in this area and describes current trends in the given topic. The practical part of the thesis consists of'. It is cited by 2. The third result is a PDF by A. Turhan from 2010, with a snippet: '• Googlebot uses an algorithmic process: computer programs determine which sites to crawl, how often, and how many pages to fetch from each site. • Googlebot's crawl process begins with a list of webpage URLs, generated from previous crawl processes and augmented with'. It is cited by 2. The fourth result is a book by A. Vegh from 2010, with a snippet: 'There can be no doubt about it: When it comes to search engines, Google is the king. According to comScore's 2008 Digital Year in Review, Google ended with 63.5% of the search market. But even more telling, Google represented 90% of all the growth in search'. It is cited by 1. The fifth result is a PDF by J. Zilincan from 2015, with a snippet: 'Abstract Search engine optimization techniques, often shortened to "SEO," should lead to first positions in organic search results. Some optimization techniques do not change over time, yet still form the basis for SEO. However, as the Internet and web design evolves'. It is cited by 1. On the left side, there are filters for 'Articles', 'Case law', and 'My library'. Below these are sorting options: 'Sort by relevance' and 'Sort by date'. There are also checkboxes for 'include patents' and 'include citations', both of which are checked. At the bottom left, there is a 'Create alert' checkbox.

**Fig 3.4: Search Page 1 (keyword : search engine optimization)**

Some of the issues needs to be concentrated on based on our search results are- citation count process of Google scholar and how Google find the patents and from where.

Basically it requires the correct bibliographic data for the cited paper. To maximize the chances that the parsers<sup>6</sup> will correctly parse the citation in the references section. The indexing guidelines need to be followed [13].It may happen that a very good article will lack the good parsing. As it is mentioned earlier Google scholar indexes patents and patent applications (patents applied for but not yet granted) from the United States Patent and Trademark Office (USPTO), and European Patent Office (EPO), which are taken from the USPTO, and the EPO databases. But it is also noticeable that there are still many organizations having patent giving authority Google Scholar doesn't consider them. Therefore, Google scholars lack many patented articles. The search results have been taken keeping Google scholar's default filtering on (including both patents and citations).

# CHAPTER 4



## 4. Findings:

We had several finding based on our study and analysis on Google Scholar Search Engine. Some of them are mentioned below:

1. The patent documents are not addressed properly. No definite hard and fast algorithm is visible for the patent document ranking compared to highly cited article ranking. The theory is also not unanimous for the importance of patent compared to scientific academic article or research.
2. Highly cited document are not always getting high rank compared to the low cited document.
3. Date of publication is another important issue. If during the search the duration or time limit is not mentioned then the precedence decision of the article does not always follow any specific order, i.e., from recent to older or the older to recent.

According to us, the highest citation count, the documents will show at first. But from the graph we can see that there have some irregularities. That is, which citation count is zero but that documents are shown in the first page at position 3 for keyword Academic Search Engine Optimization. But the citation 49 shows at the first page at first position.

But the most miserable thing is the highest citation 128 and 124 holds the position of 11 and 12 respectively at the second page. The citation count which is 137 holds the position 17 at the second page. So these are the dissimilarities which should not be happened. But still there maybe some other factor effecting on the citation count.





# CHAPTER 5



## 5. Conclusion and scope of future work:

In a word there will be no specialized result in that particular search. But the algorithm of academic search engine like Google scholar is organized in different way .It only tries to find the result for scholarly reader. That means it's main motto is to find scholarly article against in any particular search .As it can be said that academic search engine is client-specialized so it can run through the scholarly article through a quantitative process. Academic search engine helps the scientists or inventors to help their newly and most attractive innovation through most noble and innovating searching services available on the web<sup>6</sup> Parsing syntax analysis or syntactic analysis is the process of analyzing a string of symbols, either in natural language or in computer languages, conforming to the rules of a formal grammar within computational linguistics the term is used to refer to the formal analysis by a computer of a sentence or other string of words into its constituents, resulting in a parse tree showing their syntactic relation to each other, which may also contain semantic and other information.[14]

In Conclusion we can say that, there are some basic differences between general purpose search engines and academic search engines. Group of users of academic search engines are very limited. Therefore, their working mechanism isn't as efficient as a general purpose search engine. There are still scope to improve and incorporate appropriate algorithm to a academic search engine which can serve well for the academics and research community.

Through the whole thesis paper working procedure of academic search engine is tried to find out. And the most common factors are discussed briefly. There are some other factors also working behind every search.

In future, we have a plan to do an extensive analysis on the General purpose search engine to collect some advanced properties of that search engine which can be adopted in academic search engine with slight modification.

## REFERENCES:

[1]100 Time Saving Search Engines For Serious Scholars. ; By Staff Writers;  
<http://www.onlineuniversities.com/blog/2012/07/100-time-saving-search-engines-serious-scholars-revised/>

[2]<https://scholar.Google.com/>

[3]<http://academic.research.microsoft.com/>

[4]Academic Search Engine Optimization (ASEO): Optimizing Scholarly Literature for Google Scholar & Co.; Joran Beel ,Bela Gipp, Erik Wilde  
[https://docear.org/papers/Academic%20Search%20Engine%20Optimization%20\(ASEO\)%20--%20preprint.pdf](https://docear.org/papers/Academic%20Search%20Engine%20Optimization%20(ASEO)%20--%20preprint.pdf)

[5]Shari Thurow, Academic Search Engine Optimization of Web Documents for Google Scholar <https://www.asist.org/events/webinars/academic-search-engine-optimization/>

[6]Salesforce Technology Basics; Understanding the power of metadata  
[https://trailhead.salesforce.com/modules/salesforce\\_advantage\\_how\\_we\\_do\\_it/units/sa\\_how\\_power\\_of\\_metadata](https://trailhead.salesforce.com/modules/salesforce_advantage_how_we_do_it/units/sa_how_power_of_metadata)

[7]7 Common (and Dangerous) Misconceptions about SEO; Meghan Keaney Anderson;  
<https://blog.hubspot.com/marketing/seo-misconceptions-list>



[8]JöranBeel andBelaGipp. Google Scholar's Ranking Algorithm: The Impact of Citation Counts(An Empirical Study).In André Flory and Martine Collard, editors ,Proceedings of the 3<sup>rd</sup> IEEE Conference on Research Challenges in Information Science (RCIS'09), pages 439-446, Fez (Morocco),April 2009.IEEE.doi:10.1109/RCIS.2009.5089308 ISBN 978-1-4244-2865-6. Available on<http://www.sciplore.org>.

[9]JöranBeel and BelaGipp. Google Scholar's Ranking Algorithm: The Impact of Citation Counts(An Empirical Study).In André Flory and Martine Collard, editors, Proceedings of the 3<sup>rd</sup> IEEE Conference on Research Challenges in Information Science (RCIS'09), pages 439-446, Fez (Morocco),April 2009.IEEE.doi:10.1109/RCIS.2009.5089308 ISBN 978-1-4244-2865-6. Available on<http://www.sciplore.org>.

[10] h-index ;[https://en.wikipedia.org/wiki/H-index#cite\\_note-1](https://en.wikipedia.org/wiki/H-index#cite_note-1)

[11]Joseph G. Hadzima, Jr. The Importance of Patents: It Pays To Know Patents Rules. <https://web.mit.edu/e-dub/hadzima/pdf/the-importance-of-patents.pdf>.

[12]AlirezaNoruzi, Mohammadhiwa Abdelkhoda; Google Patents: The global patent search engine. [https:// www.webology.org/2014/v11n1/a122.pdf](https://www.webology.org/2014/v11n1/a122.pdf).

[13]Inclusion Guidelines for Webmasters;<https://scholar.Google.com/intl/en-us/scholar/inclusion.html#indexing>.

[14]Parsing; <https://en.wikipedia.org/wiki/Parsing>.