

# Search Engine Optimization: A Game of Page Ranking

**AanchalKakkar**

ASET, Amity University, Noida  
**Email Id:** kakkar.anchal@gmail.com

**RanaMajumdar**

ASET, Amity University, Noida  
**Email Id:** rmajumdar@amity.edu

**Arvind Kumar**

ASET, Amity University, Noida  
**Email Id:** arvind.jki@gmail.com

*Abstract: Man made machine has made the world dynamic as the universe of digitalization has arrived making an accounts unimaginable. Anything taught to be is unthinkable and can be turned to conceivable inside our own particular safe place keeping the data secure. This has been conceivable with the assistance of web crawler. The paper clarifies Google's new calculation HUMMINGBIRD which thought to overwhelm SEO systems yet luckily SEO does not have a short life compass. This paper tries to explain the comparison of quality content of different google search algorithms. It also through the light on SEO strategies. The paper concludes with "SEO can never have a short life compass".*

**Keywords – SEO, hummingbird, google, internet, traffic, social networking, page rank, page speed**

## I.INTRODUCTION

Presentation of online networking has prompted the help of online business showcasing. None of us can consider carrying on with our life exclusively, none of us can turn into our own particular closest companions, offering sensitivity to our self or to get by without comprehending what's going on the planet. An enormous test confronting society today is the means by which to structure system and organizations together among individuals, on both nearby and worldwide levels. Today everybody accepts to unite with individuals hearts and manufacture new grassroots systems. Trying endeavors to converse with and chat with others is the initial phase in changing oneself from a shut to an open person which makes companionship a lifetime treasure. What's more today INTERNET has made everything conceivable. One can unite with any corner of the world, and web has planted the seeds of inquiry, social networking and substance as strategy.

Web search tool, a best apparatus to empower clients to spot any data on overall web and to build a positioning in internet searcher an idea known as SEO entered. We compose our examination after fitting investigation of SEO method and its execution through Page Ranking instrument.

SEO remains for website improvement, a procedure to assemble a decent substance, understanding the ways individuals may look for it. In expansion SEO incorporates the methodology of decisive word research.

There are numerous profits of SEO which incorporates:

- i. promotes business at monetary scale.
- ii. Increase movement and site fame.
- iii. Returns on speculation to perky contender site.
- iv. Cost viability to give required data just.

Look calculation is the heart of web search tool. The emotional change performed by google was in 2001.google's new approach to move past essential words and underscore on "substances" prompts improvement of hummingbird. A greatest affirmation was overhauled to its internet searcher in over 10 years, which it named HUMMINGBIRD.

Hummingbird: A real revamp of Google's whole new calculation is composed extraordinarily to convey particular responses for complex questions. The Hummingbird calculation influenced 90% of hunts and was reported nearby an alternate real change for clients: Google's scrambling of essential word information. The move of encoding decisive word information will basically render pointless the "Natural Traffic Sources" Google Analytics information utilized by numerous SEO's [1].even humors spread around that google has prompted the end of SEO and our paper will incorporate the finish of the gossipy tidbits and will come close which subject to be embraced by business site to draw in more of the clients.

Google calculation upgrades will dependably help the objective of showing quality substance which is adjusted to the data needs of searchers. Distinctive calculations have been utilized by Google where disadvantage of every calculation prompts the improvement of new calculation.

TABLE I. COMPARISON OF QUALITY CONTENT OF DIFFERENT SEARCHERS [2]

Google Penguin	Google Panda	Google Hummingbird
It expects to diminish the perceivability for the destinations which takes part in the connection plans and/or spammy practices, for example, catchphrase stuffing	It goes for decreasing the perceivability for low quality substance locales, for example, those that are excessively substantial with the publicizing	It is focused around the semantics look which concentrates on client aim.  It goes for enhancing query items for customary questions, for example, questions.

It expects to diminish the deceivability for the locales which takes part in the connection plans and/or spams practices, for example, decisive word stuffing it goes for diminishing the deceivability for low quality substance destinations, for example, those that are excessively substantial with the advertising it is focused around the semantics seek which concentrates on client goal. It goes for enhancing indexed lists for ordinary questions, for example, questions.

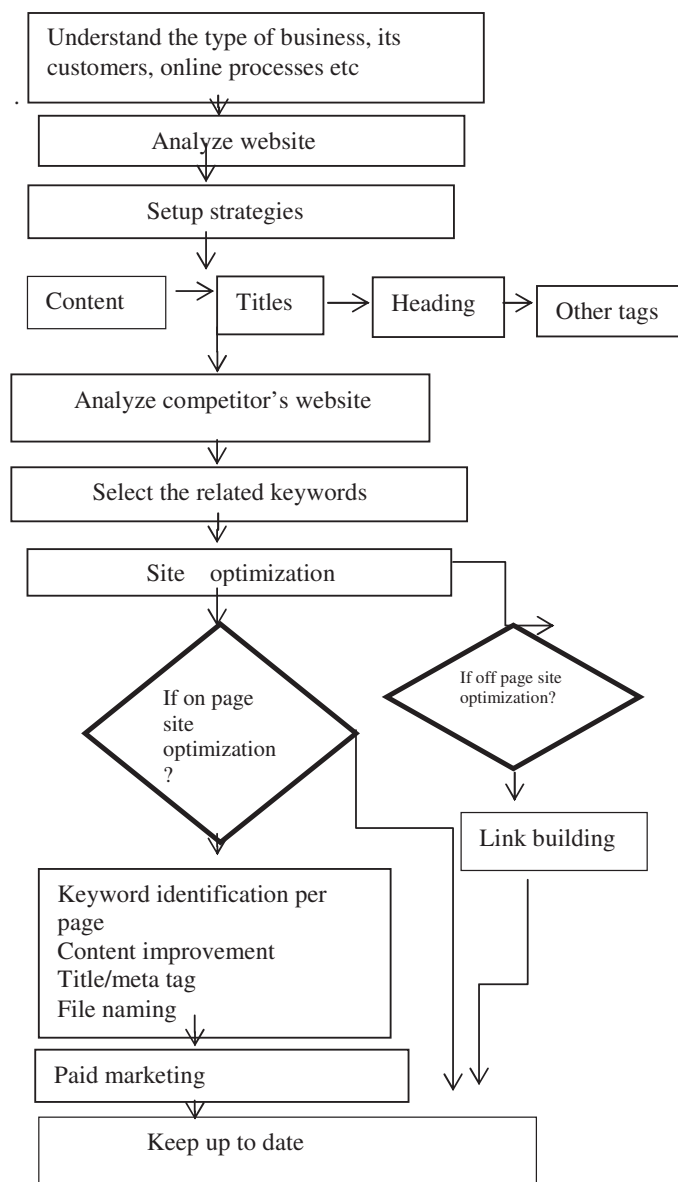


Fig. 1. flowchart depicting the workflow of SEO.

With the understanding of different SEO apparatuses including Page velocity, code to content degree checker, substance

matching we delineate the working of web index into flowchart delineating how each one stage assumes a paramount part for the coordinators to draw in more customers towards his business. It is vital for the business to legitimately dissect and set up the different methods so as to win the trust of their customers. The accompanying table outlines the procedures as flowchart that can make the understanding of the work process of SEO less demanding.

#### A. Search Engine Optimization

Some substance [3] creators arrange their destinations not just to pull in and keep up the consideration of human crowds additionally oblige and even exploit web crawlers with their positioning guidelines to situate website improvement to an expert speciality. Some organizations of internet searchers distribute SEO guidance to the webmasters to advance the idea of "white cap" real, gathering of people administered, where results endures long time, advancing availability and is a summation of substance for clients not for web indexes. To facilitate the exertion of SEO practices to precisely read sites to dishearten webmasters and web advertisers from turning to "dark cap" misleading, where result can at times be expected to boycott the webpage incidentally or forever and uses concealed content which can either be hued like the foundation to look for the positioning framework.

#### B. Survival Of The Fittest

Today the opposition has made everybody to be similar to running water to manage in present. Indeed the page rank of the site builds the quantities of guests. Essentially SEO has been ordered in two separate classifications which help to keep up the procedures for SEO so as to make the correspondence channel occupied.

In 2010 [4] it was affirmed that Google and Bing consider connections, notoriety and power to connection writer on social networking including Twitter and Facebook which incorporates their positioning variable. A future may arrive where social signs will get more consideration from internet searchers prompting the idea of SMO, pursuit media streamlining "upgrade a webpage which can be effectively connected and is exceedingly unmistakable in online networking hunt on custom web indexes where much of the time incorporate significant posts on websites, podcasts and sites, SMO must be considered". Indeed SMO methods can be significantly considered a piece of SEO.

#### C. Seo Strategy

Basically search engine is based on the behavior of the casual users where they just put up the simple query on most popular topics [5]. The availability of popular queries and blog post on popular topics leads to the better understanding of what web users want to know about and what they are talking about. The opinion obtained with the interactions from queries, blogs, news and various social media is detected in a user-oriented

manner where [6] recently google Flu trends released an application which can estimate the flu epidemics based on queries which has been received by massive web users. Basically different strategies are applied on search engine optimization including:

- i. Social media
- ii. Mobile SEO
- iii. Structured data
- iv. Content optimization
- v. Conversational content
- vi. Driving relevant

Google tried to undertake social media authority signals to make primary communication channel with customers and focus on the following areas:

- i. Google authorship
- ii. Article publisher markup

Google algorithm “hummingbird” and search engine optimization algorithm have different focus area which is depicted in the below table:

TABLE II. SUMMARIZATION OF FOCUS AREA [7]

Focus Of Hummingbird	Focus Of Seo
Google used SSL encryption search results for both offline and online users	It focused on various sorting technology
Search engine will become much more human friendly.	Less human friendly.
Increase of smartphones, android and tablets has led to emergence of mobile SEO	Website mobile performance will be the issue matter of SEO ranking

## II. LITERATURE REVIEW

In 2007 [8] at Indiana college, Menczer was making another sort of internet searcher which was focused around web metadata-catchphrase. He was taking a shot at Give join (www.givealink.org) which is truly like del.icio.us (http://del.icio.us). this long range informal communication webpage let the enlisted clients to impart their bookmarks online and can without much of a stretch label them with significantly more Meta data. Indeed a news aggregator Digg (www.digg.com) lets the clients vote in favor of sites which appears imperative to them. Dealing with long range interpersonal communication Menczer's group connected machine learning calculation when guests transfer their bookmarks. Applying machine learning calculation they guide out connections between the transferred bookmarks which help them in making a positioning plan for web search tool.

### A. Page rank

To check the Page need or to check the page has been recently overhauled gives the thought of page positioning. Page rank assumes an exceptionally pivotal part in site design improvement. The principal page in google turns into the most productive page for the client to get to it and get the learning

and accomplish the work. Higher the page rank higher the need of clients to get to it [9]

Page rank calculation is singularly focused around connection structure which is characteristic on the web where each one page is thought to be voting the pages to which it is joined in this manner reflecting the aggregate number of votes out of which imperative vote is reflected. This issue is scientifically detailed as discovering the Eigen vector relating to biggest Eigen estimation of certain stochastic lattice connected with web structure.

Since the web is said to be made out of in excess of 9 billion pages and still the size is developing the page rank processed by google with the assistance of crawlers which naturally search the web. Page rank [10] calculation gives the critical part of pages for web crawler, for example, Google which allot to each one site page a measure of its criticalness and the ubiquity singularly focused around the connection structure of the web which has prompted the test in executing the calculation size of the web where files is gathered at Google is in excess of 10 billion bringing on genuine issues for processing.

Using the SEO tool the page rank of given URL can be found as follow:



Fig. 2. Page rank [11]

A formulae [12] is used for calculating the ranking of a page

$$PR(A) = \left( \frac{1-d}{n} + d \sum_{i=1}^n \frac{PR(T_i)}{C(T_i)} \right) \text{ -----equation (1),}$$

Where every variable defined as :

PR (A) = values of A's page rank

C (Ti) = number of links from Page Ti.

D = Damppling coefficient values between 0 and 1

N=number of pages on internet.

Henshaw [13] stressed that the heart of search engine is search algorithms which analyzes and make senses of billions of pages to provide most relevant search. Google's hummingbird algorithm affects more than 90% of search result.

Marketing [14] has become more sophisticated with powerful marketing channels. The progression of economy growth lead to reach to a common understanding of marketing trends and activities lauding to vast changes to reach to the customer's behavior and pattern.

No internet marketing company [15] can discount the power of google where PR newswire noted that SEO owns more than two third of market share.

*B. Focused Area*

Page rank is the major focused area which depicts that page rank plays a very vital role to attract more clients towards the site. Iterative algorithm has been used to calculate the page rank corresponding to the normalized link matrix of the web. Search engine algorithm is implemented on a MATLAB to check the effect of iteration on page rank.

The page rank can be calculated using below formulae with various dependencies factors:

$$PageRank(A) = (1-d) + d \left( \frac{PageRank(T_1)}{C(T_1)} + \dots + \frac{PageRank(T_n)}{C(T_n)} \right)$$

-----equation (2)

To determine the ranking of webpage A, we take into account the rankings of every other page on the web.

**Page Rank (Tn)** –Each page represents its occurrence in the web from first page to the nth page.

**C (Tn)** –It represents the count or number of link of the page associated with it.

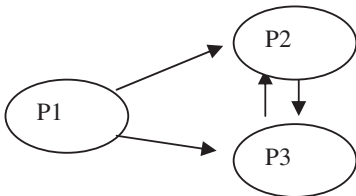
**PR (Tn) /C (Tn)** –It leads to the count of vote of page A that represents that page A has a backlink from page n.

**d** – It is a damped down value with a factor of 0.85 which prevents the influence of other pages.

**(1-d)** –Limit of a page rank lies between 0-10.(1-d) will give (1-0.85=0.15) representing the minimum value of a page rank even if page has no links attached to it.

III.CASE STUDY

On a proper understanding of Impact of iteration on page rank we have formulize our work using a case study. A small observation has been noted down in order to estimate the page rank values at various intervals.



Suppose three different pages are linked together such as:

$$A_{ij} = 1/n_j \text{ if } j \in L_i$$

$$0 \text{ otherwise}$$

So matrix can be calculated as:

0	1/2	1/3
0	0	1
0	1	0

Modified matrix can be constructed as:

$$M = (1-m)A + (m/n)S$$

Where m chosen be 0.15 and S=1

0.0375	0.4625	0.3208
--------	--------	--------

0.0375	0.0375	0.8875
0.0375	0.8875	0

Thus final output obtained is [0.0375 0.425 0.116]

*A. Result Obtained*

The result shows that page 2 has the highest value amongst them because it is linked to large number of pages and page 1 has the smallest value since it is linked to small number of pages.

Iterative algorithm has been used to calculate the page rank corresponding to the normalized link matrix of the web.

Search engine algorithm is implemented on a MATLAB to check the effect of iteration on page rank. A small observation has been noted down in order to estimate the page rank values at various intervals. We have implemented the iterative algorithm on two pages to compute the effect of iterations on the page rank. These cases compute that how navigation of a page affects its ranking.

**CASE : estimating PR=0**

Number of iterations	PR(1)	PR(2)
1	.15	.2775
2	.385875	.47799375
3	.556294875	.6228504844
4	.6794229117	.727509475
5	.7683830537	.8031255957
6	.8326567563	.8577582429
7	.8790945064	.8972303305
8	.9126457809	.9257489138
9	.9368865765	.9463535902
10	.9544005517	.9612404689
11	.9670543986	.9719962388
12	.976196803	.9797672825
13	.9828021901	.9853818616
14	.9875745824	.989438395
15	.9910226358	.9923692404
16	.9935138543	.9944867762
17	.9953137598	.9960166958
18	.9966141914	.9971220627
19	.9975537533	.9979206903
67	.999999997	1.14

Suppose the webpage 1 and webpage 2 are newly developed pages which have never been navigated by any of the user so applying the iterative algorithm it has been estimated that around 67 iterations would be required for page 2 to make it to stand at the ranking of 1.page 1 still requires more iterations. The PR of page is calculated by google spider once it come in contact with it which has no idea about the content or size of the page or the language and the text used in the link.URL strips off to the parent page when present on the toolbar.

- 1) The normalized probability distribution of page rank will be 1.0
- 2) When number of pages includes and when we start page with 1, the little iteration is required to converge to a suitable result.



- 3) No negative value can be estimated.
- 4) Each time an algorithm is executed it gets closer to the final value.

#### IV.METHODOLOGY ADOPTED FOR THE CALCULATION OF PAGE RANK

The PR of page is calculated by google spider once it come in contact with it which has no idea about the content or size of the page or the language and the text used in the link.URL strips off to the parent page when present on the toolbar.

With the proper understanding of page rank under different cases we have summarized the result as follow:

- i. The normalized probability distribution of page rank will be 1.0
- ii. When number of pages includes and when we start page with 1, the little iterations required to converge to a suitable result.
- iii. No negative value can be estimated.
- iv. Each time an algorithm is executed it gets closer to the final value.

#### V.CONCLUSION

Individuals are joined with one another through social networking prompting the rare development in informal organization. Communists frequently inspect contradicting model in which society decides the configuration and acknowledgement of innovation.

It has been assessed that disregarding the social connection of reception can prompt disappointment actually when development is required by target group however Hays [16] registers that "states with more noteworthy societal issues reacts with the weaker laws" as dubious nature of an issue affect the significant arrangements. A ton of examination is centered around web crawler innovations which today have turned into an indispensable piece of life featuring from a kid to old ones. Indeed one can do shopping sitting under a top with cooled room.Withgoogle hummingbird, SEO ought to concentrate on client engagement as opposed to essential word information.

As Google's Distinguished Engineer Matt Cutts has said: "Succeeding in SEO will be the same as its generally been in case you're doing it right – give the clients an extraordinary experience [17].although dispatch of hummingbird calculation didn't surpass much however Google's dispatch of learning diagram is truly an impetus to comprehend semantic web and subsequently SEO can never have a short life compass. Past information has uncovered that weakness of inquiry calculation was SEO strategies which were not ready to distinguish the "dark cap" systems and different spam sites were positioning top falsely. While perceiving different interpersonal interaction instruments we reason that that SEO has a high life compass and this review supplements considerable ebb and flow take a shot at idea discovering that website streamlining is subject to calculation heading clients to use web seek. Accordingly we

have reasoned that no negative worth can be assessed for the page rank.

#### REFERENCES

- [1] C. Lin and R. Yazdanifard, "How Google's New Algorithm, Hummingbird, Promotes Content and Inbound Marketing," American Journal of Industrial and Business Management, Vol. 4 No. 1, 2014, pp. 51-57.
- [2] [www.slideshare.net/ChrisSchweppe/gsi-hummingbird-pov-20131003-presofinal](http://www.slideshare.net/ChrisSchweppe/gsi-hummingbird-pov-20131003-presofinal)
- [3] John B.Killoran-"how to use search engine optimization techniques to increase website visibility"IEEE transaction on professional communication vol 56 no.1 march 2013
- [4] SaeidSaberietal "Past, Present and Future of Search Engine Optimization" IEEE ISBN: ISBN: 978-0-9891305-2-3 ©2013 SDIWC
- [5] Aixin sun and Meishan Hu-"query guided event detection from news and blog streams"IEEE transaction on systems,MAN and cybermetrics Part A:Systems and humans vol 41 no.5 sept 2011
- [6] J.Ginsberg,M.HMohebbi,R.SPatel,L.Brammer,M.SSmolinski and L.Brilliant(2009 feb) detecting influwnza epidemics using search engine query data nature[online].457(7232)pp 1012-1014.Available:http://dx.doi.org/10.1038/nature 07634
- [7] [www.slideshare.net/jhenrey1992/seo-seminar-29202765?qid=0dbfb271-5cb2-4c52-a67a-](http://www.slideshare.net/jhenrey1992/seo-seminar-29202765?qid=0dbfb271-5cb2-4c52-a67a-)
- [8] Rubin Landau "building better search engines"Propogation society July/august 2007 IEEE Antennas and by ISN-1521-9615/07 copublished by IEEE CS and AIP
- [9] Hideaki Ishii and Roberto Tempo-"Distributed Randomized algorithms for page rank computation"-IEEE transaction on automatic control vol 55 no.9,sept 2010.
- [10] Hideaki Ishii,Roberto Tempo and Er-Wei-Bai "A web aggregation approach for distributed randomized Page rank algorithm"-IEEE transaction on Automatic control vol 57,no.11 November 2012
- [11] PR Newswire, "Internet Marketing Company Cyberset,"Adapts to Google's New Algorithm, Hummingbird," 2013.
- [12] Rakesh Kumar Shivasaini"a study on SEO monitoring system based on corporate website development"ijcseit vol1 June 2011
- [13] Renyuanwang,Yanchen,TaoyingLi,Ying Ying Yui"the optimization of search engine ranking technology based on grey system",2013 IEEE,ICCSIS
- [14] J. Henshaw, "What Google's Hummingbird Update Means for Content Marketers," 2013.
- [15] ChauntelleOng Yi Lin, RashadYazdanifard <http://smallbusiness.yahoo.com/advisor/google-hummingbird-means-content-marketers-143119302.html>(American Journal of Industrial and Business Management, 2014, 4, 51-57 Published Online January 2014
- [16] leoA.meyerovich,Arielrabkin,"socio-PLT: principles for programming language adoption" Oct 26,2012
- [17] Webrefrence"(<http://searchenginewatch.com/article/2300838/Google-Keyword-Not-Provided-How-to-Move-Forward>