International Journal of Science, Engineering and Technology Research (IJSETR), Volume 3, Issue 4, April 2014

A Survey to Enhance the Page Ranking Algorithm

S.Sowbaranika, B.Thirunavukarasu, S.Stewart Kirubakaran, Dr.S.Karthik

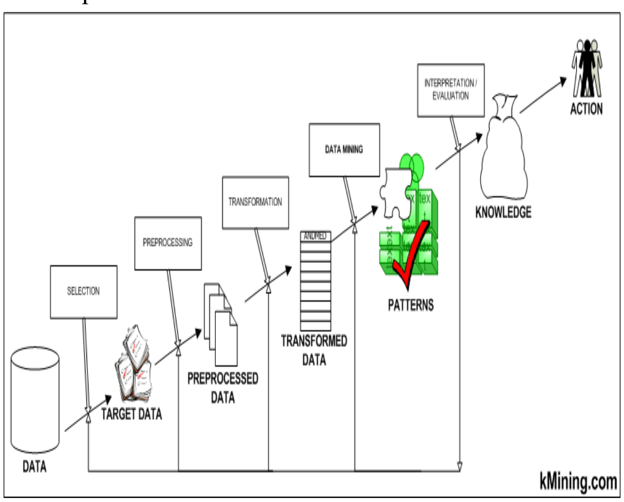
***Abstract*—We live in a computer era. Internet is part of our everyday lives and information is only a click away. Communicating across the web has become an integral part of everyday life. This communication is enabled in part by scientific studies of the structure of the web. We consider a simple model, known as the random surfer model. We consider the web to be a fixed set of pages, with each page containing a fixed set of hyperlinks, and each link a reference to some other page. We study what happens to a person (the random surfer) who randomly moves from page to page, either by typing a page name into the address bar or by clicking a link on the current page.**

***Index Terms* — Data Mining, Page Ranking Algorithm, Random surfer model, The PageRank display of the google toolbar, Google directory, characteristics, iterative components**

.

I. INTRODUCTION

Data mining is a burgeoning sub-field in computer science wherein search, collection and retrieval of data is done from large volumes of databases. Useful patterns



and consumable data are segregated from those databases and are kept for future use in data warehouses.

Sowbaranika S,*Computer Science and Engineering Department, SNS College of Technology, ., Coimbatore, INDIA.*

Thirunavukarasu B, *Computer Science and Engineering Department, SNS College of Technology. Coimbatore, INDIA.*

Stewart Kirubakaran S, Assistant Profressor,*Computer Science and Engineering Department, SNS College of Technology, Coimbatore, INDIA.*

Dr Karthik, Professor and Dean, *Computer Science and Engineering Department, SNS College of Technology, Coimbatore, INDIA.*

Refined algorithms are used for extraction of data from the databases. The major use of data mining is to improve the standards of any database management systems. The production process could be enhanced with minimal time consumption and the possible outcomes of a system could be predicted efficiently so that the system would become more tolerant to miscellaneous situations it faces. Data mining encompasses the ideas of wide variety of fields such as machine vision technology, image processing, artificial intelligence, neural networks etc.Data mining has wide variety of applications in business, banking sector, transaction management, information repositories, data analysis, retail manufacturing unit, telecommunication, marketing, software engineering, data visualization etc.,

There are many types of data mining. They are sequential patterns, sequence similarity, clustering, classification; association rules etc., one of the greatest applications of data mining is web mining. Here patterns and data are extracted from the web and the services provided by the web applications can be immensely increased. The process involves fetching of data from many users and from many websites, analysis of the data amassed to select needed information from it, production of many useful and innovative applications and services. Thus web mining-an intensified version of data mining could aggregate more quantity of data instantly
and could even render a revolution in the web world.

A. Manipulation of page rank

Reputation and ranking systems are an essential part of web search and ecommerce. The general idea is that the reputation of one participant is determined by the endorsements of others; for example, one web page endorses another by linking to it. However, not all participants are honorable, so manipulation-resistance is an important consideration. Link analysis algorithms for web search first appeared in the late 1990s. Notable examples are HITS and Page Rank [7]; these algorithms use the global hyperlink structure of the web to determine the importance of individual pages. Each uses notions of endorsements among pages, so we also refer to them as link-based reputation systems. Our work focuses on Page Rank, an influential link-analysis algorithm developed by Sergey Brin and Lawrence Page to rank web pages in the original Google search engine [7]. Page Rank is a success story: it is still today an important ingredient in search technology, and it has been exported to many other domains. However, because Page Rank is so successful, it is also a popular target for manipulation in the form of link spam, where web authors manipulate hyperlinks to boost their own reputation. There is a strong economic incentive for pages to rank highly in search results, and Page Rank is a well-known and simple-to-understand part of a search engine. Moreover, despite the global nature of the Page Rank algorithm, it is relatively easy for a page to boost its Page-Rank by changing only out links. This is an undesirable property in a reputation system: one would like the reputation of a page to determine by the actions of other participants.

B. An outlook about page rank

It is instructive to take a closer look at the Page Rank to see where things go wrong. Let G = (V, E) be a directed graph (e.g., the web). Page Rank assigns the score _ (v) to node v, where _ is defined to be the stationary distribution of
a random walk on G, giving the pleasing interpretation that the score of page v is the fraction of time a web surfer spends there if she randomly follows links forever. For technical reasons, the random walk is modified to restart in each step with probability _, jumping to a page chosen at random, either uniformly or from a pre-specified distribution. Informally, this allows the random walk to escape "dangling" portions of the web graph that do not connect back to main portion. Mathematically, it ensures that _ exists, is unique, and efficient to compute. A well-known fact about Markov chains is that the PR(Tn)/C(Tn))
where,

- PR(A) – Page Rank of page A
- PR (Ti) – Page Rank of pages Ti which link to page A
- C (Ti) - number of outbound links on page Ti
- d - damping factor which can be set between 0 and 1

A simple way of representing the formula is, (d=0.85)
Page Rank (PR) = 0.15 + 0.85 * (a share of the Page Rank of every page that links to it)

D. Disadvantages

1. More responsibility to maintain if higher page rank is sought,

2. Face more spammers on your commenting system, because everybody wants back links from higher Page Rank and good authority website,

Page Rank is good but sometimes it leaves bad thoughts in the mind of people because if you'll search on Google most of the website is having low page rank i.e. 2 or 3 or sometimes 0 and it is coming on top of the page of Google.

Some organization or enterprises are very much interested in increasing their page rank instead of growing their business. It is a fact that with increase in page rank your business will not grow.

You should be concentrating more on the future growth

expected return time of v — the number of steps it takes a random walk starting at v to return to v — is equal to 1/_ (v) [1]. A heuristic argument for this equivalence is that a walk that returns to v every r steps on average should spend 1/r of all time steps there.

C. Manipulation by outlinks

The intuition of the manipulation strategy follows easily from the interpretation of Page Rank as the inverse of expected return time. If node v wishes to minimize the expected return time, it should link only to nodes from which a random walk will return to v very quickly, in expectation. By partnering with just one other node to form a 2-cycle with no other out links, v ensures a return in two steps — the minimum possible assuming self-loops is ignored — unless the walk jumps first.

II. PAGE RANKING ALGORITHM

Page ranking is a methodology to sort out the importance of pages that are found on the web. Many algorithms and procedures have been devised to discover the rank of the pages. It is a numerical value. When one page makes a link to an another page, a vote is cast to the latter page Rank of a page is decided based upon the number of votes casted to it or more handily based upon the usage of the page.

The original page ranking algorithm was put forward by Larry Page and Sergey Brin
PR(A) = (1-d) + d (PR(T1)/C(T1) + ... +
of your company instead of your page rank because google updates page rank with a span of three to four months, which is usually seen.

Well the last update which google does on page rank is on the end of 2008 where most of the website whose page rank is 0 or 1 increased to 3 or 4 and others where having more than 5 or 6 decrease to 3 or 4.

E. Applications:

Well, nobody understand the logic of Google behind giving the rank to a website or what are they considering the most which according to them is best.
Computer forensics is the traces collection and processing, which is the offender to remain in the computer or network system, and as a legally binding evidence in the proceedings available to the court, so that suspects would be brought to justice. It mainly includes data protection, data collection, data analysis, the evidence presented in such processes. the data analysis is the key to computer forensics. It is faced with a problem how to extract data from a mass suspected data and core criminal evidence collection. This article applies Page Rank algorithm in computer forensics and we can obtain the crime evidence set.

It is a well-known problem that intrusion detection systems overload their human operators by triggering thousands of alarms per day. This paper presents a new approach for handling intrusion detection alarms more efficiently. Central to this approach is the notion that each alarm occurs for a reason, which is referred to as the alarm's root causes. This paper observes that a few dozens of rather persistent root causes generally account for over 90% of the alarms that an intrusion detection system triggers. Therefore, we argue that alarms should be handled by identifying and removing the most predominant and persistent root causes. To make this paradigm practicable, we propose a novel alarm-clustering method that supports the human analyst in identifying root causes. We present experiments with real-world intrusion detection alarms to show how alarm clustering helped us identify root causes. Moreover, we show that the alarm load decreases quite substantially if the identified root causes are eliminated so that they can no longer trigger alarms in the future.

III. THE RANDOM SURFER MODEL

Lawrence Page and Sergey Brin give a very simple intuitive justification for the PageRank algorithm. They consider PageRank as a model of user behaviour, where a surfer clicks on links at random with no regard towards content.

The random surfer visits a web page with a certain probability which derives from the page's PageRank. The probability that the random surfer clicks on one link is solely given by the number of links on that page. This is why one page's PageRank is not completely passed on to a page it links to, but is divided by the number of links on the page.

So, the probability for the random surfer reaching one page is the sum of probabilities for the random surfer following links to this page. Now, this probability is reduced by the damping factor d. The justification within the Random Surfer Model, therefore, is that the surfer does not click on an infinite number of links, but gets bored sometimes and jumps to another page at random.

The probability for the random surfer not stopping to click on links is given by the damping factor d, which is, depending on the degree of probability therefore, set between 0 and 1. The higher d is, the more likely will the random surfer keep clicking links. Since the surfer jumps to another page at random after he stopped clicking links, the probability therefore is implemented as a constant (1-d) into the algorithm. Regardless of inbound links, the probability for the random surfer jumping to a page is always (1-d), so a page has always a minimum PageRank.

A. A different notation of the pagerank algorithm

Lawrence Page and Sergey Brin have published two different versions of their PageRank algorithm in different papers. In the second version of the algorithm, the PageRank of page A is given as

$$PR(A) = (1-d) / N + d\,(PR(T1)/C(T1) + ... + PR(Tn)/C(Tn))$$

Where ‘N’ is the total number of all pages on the web. The second version of the algorithm, indeed, does not differ fundamentally from the first one. Regarding the Random Surfer Model, the second version's PageRank of a page is the actual probability for a surfer reaching that page after clicking on many links. The PageRanks then form a probability distribution over web pages, so the sum of all pages' PageRanks will be one.

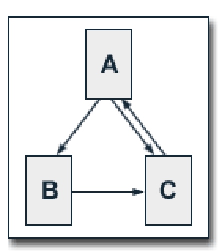
Contrary, in the first version of the algorithm the probability for the random surfer reaching a page is weighted by the total number of web pages. So, in this version PageRank is an expected value for the random surfer visiting a page, when he restarts this procedure as often as the web has pages. If the web had 100 pages and a page had a PageRank value of 2, the random surfer would reach that page in an average twice if he restarts 100 times.

As mentioned above, the two versions of the algorithm do not differ fundamentally from each other. A PageRank which has been calculated by using the second version of the algorithm has to be multiplied by the total number of web pages to get the according PageRank that would have been calculated by using the first version. Even Page and Brin mixed up the two algorithm versions in their most popular paper "The Anatomy of a Large-Scale Hyper textual Web Search Engine", where they claim the first version of the algorithm to form a probability distribution over web pages with the sum of all pages' PageRanks being one.

In the following, we will use the first version of the algorithm. The reason is that PageRank calculations by means of this algorithm are easier to compute, because we can disregard the total number of web pages.

IV. THE CHARACTERISTICS OF PAGERANK

The characteristics of PageRank shall be illustrated by a small example.



We regard a small web consisting of three pages A, B and C, whereby page A links to the pages B and C, page B links to page C and page C links to page A. According to Page and Brin, the damping factor d is usually set to 0.85, but to keep the calculation simple we set it to 0.5. The exact value of the damping factor d admittedly has effects on PageRank, but it does not influence the fundamental principles of PageRank. So, we get the following equations for the PageRank calculation:

PR(A) = 0.5 + 0.5 PR(C)

PR(B) = 0.5 + 0.5 (PR(A) / 2)

PR(C) = 0.5 + 0.5 (PR(A) / 2 + PR(B))

These equations can easily be solved. We get the following PageRank values for the single pages:

PR(A) = 14/13 = 1.07692308

PR(B) = 10/13 = 0.76923077

PR(C) = 15/13 = 1.15384615

It is obvious that the sum of all pages' PageRanks is 3 and thus equals the total number of web pages. As shown above this is not a specific result for our simple example.

For our simple three-page example it is easy to solve the according equation system to determine PageRank values. In practice, the web consists of billions of documents and it is not possible to find a solution by inspection.

A. The iterative computation of pagerank

Because of the size of the actual web, the Google search engine uses an approximate, iterative computation of PageRank values. This means that each page is assigned an initial starting value and the PageRanks of all pages are then calculated in several computation circles based on the equations determined by the PageRank algorithm. The iterative calculation shall again be illustrated by our three-page example, whereby each page is assigned a starting PageRank value of 1.

Iteration	PR(A)	PR(B)	PR(C)
0	1	1	1
1	1	0.75	1.125
2	1.0625	0.765625	1.1484375
3	1.07421875	0.76855469	1.15283203
4	1.07641602	0.76910400	1.15365601
5	1.07682800	0.76920700	1.15381050
6	1.07690525	0.76922631	1.15383947
7	1.07691973	0.76922993	1.15384490
8	1.07692245	0.76923061	1.15384592
9	1.07692296	0.76923074	1.15384611
10	1.07692305	0.76923076	1.15384615
11	1.07692307	0.76923077	1.15384615
12	1.07692308	0.76923077	1.15384615

We see that we get a good approximation of the real PageRank values after only a few iterations. According to publications of Lawrence Page and Sergey Brin, about 100 iterations are necessary to get a good approximation of the PageRank values of the whole web.

Also, by means of the iterative calculation, the sum of all pages' PageRanks still converges to the total number of web pages. So the average PageRank of a web page is 1. The minimum PageRank of a page is given by (1-d). Therefore, there is a maximum PageRank for a page which is given by dN+ (1-d), where N is total number of web pages. This maximum can theoretically occur, if all web pages solely link to one page, and this page also solely links to itself.

B. The implementation of pagerank in the google search engine

Regarding the implementation of PageRank, first of all, it is important how PageRank is integrated into the general ranking of web pages by the Google search engine. The proceedings have been described by Lawrencec Page and Sergey Brin in several publications. Initially, the ranking of web pages by the Google search engine was determined by three factors:

- Page specific factors
- Anchor text of inbound links
- PageRank

Page specific factors are, besides the body text, for instance the content of the title tag or the URL of the document. It is more than likely that since the publications of Page and Brin more factors have joined the ranking methods of the Google search engine. But this shall not be of interest here.

In order to provide search results, Google computes an IR score out of page specific factors and the anchor text of inbound links of a page, which is weighted by position and accentuation of the search term within the document. This way the relevance of a document for a query is determined. The IR-score is then combined with PageRank as an indicator for the general importance of the page. To combine the IR score with PageRank the two values are multiplicities. It is obvious that they cannot be added, since otherwise pages with a very high PageRank would rank high in search results even if the page is not related to the search query.

Especially for queries consisting of two or more search terms, there is a far bigger influence of the content related ranking criteria, whereas the impact of PageRank is mainly visible for unspecific single word queries. If webmasters target search phrases of two or more words it is possible for them to achieve better rankings than pages with high PageRank by means of classical search engine optimisation.

If pages are optimised for highly competitive search terms, it is essential for good rankings to have a high PageRank, even if a page is well optimised in terms of classical search engine optimisation. The reason therefore is that the increase of IR scores deminishes the more often the keyword occurs within the document or the anchor texts of inbound links to avoid spam by extensive keyword repetition. Thereby, the potentialities of classical search engine optimisation are limited and PageRank becomes the decisive factor in highly competitive areas.

C. The pagerank display of the google toolbar

PageRank became widely known by the PageRank display of the Google Toolbar. The Google Toolbar is a browser plug-in for Microsoft Internet Explorer which can be downloaded from the Google web site. The Google Toolbar provides some features for searching Google more comfortably.

The Google Toolbar displays PageRank on a scale from 0 to 10. First of all, the PageRank of an actually visited page can be estimated by the width of the green bar within the display. If the user holds his mouse over the display, the Toolbar also shows the PageRank value. Caution: The PageRank display is one of the advanced features of the Google Toolbar. And if those advanced features are enabled, Google collects usage data. Additionally, the Toolbar is self-updating and the user is not informed about updates. So, Google has access to the user's hard drive.

If we take into account that PageRank can theoretically have a maximum value of up to dN+(1-d), where N is the total number of web pages and d is usually set to 0.85, PageRank has to be scaled for the display on the Google Toolbar. It is generally assumed that the scalation is not linearly but logarithmically. At a damping factor of 0.85 and, therefore, a minimum PageRank of 0.15 and at an assumed logarithmical basis of 6 we get a scalation as follows:

Toolbar-PR	Tatsächlicher PR	
0/10	0.15	- 0.9
1/10	0.9	- 5.4
2/10	5.4	- 32.4
3/10	32.4	- 194.4
4/10	194.4	- 1,166.4
5/10	1,166.4	- 6,998.4
6/10	6,998.4	- 41,990.4
7/10	41,990.4	- 251,942.4
8/10	251,942.4	- 1,511,654.4
9/10	1,511,654.4	- 9,069,926.4
10/10	9,069,926.4	- 0.85 × N + 0.15

It is uncertain if in fact a logarithmical scalation in a strictly mathematical sense takes place. There is likely a manual scalation which follows a logarithmical scheme, so that Google has control over the number of pages within the single Toolbar PageRank ranges. The logarithmical basis for this scheme should be between 6 and 7, which can for instance be rudimentary deduced from the number of inbound links of pages with a high Toolbar PageRank from pages with a Toolbar PageRank higher than 4, which are shown by Google using the link command.

D. The toolbar's pagerank files

Even webmasters who do not want to use the Google Toolbar or the Internet Explorer permanently for security and privacy concerns have the possibility to check the PageRank values of their pages. Google submits PageRank values in simple text files to the Toolbar. In former times, this happened via XML. The switch to text files occured in August 2002.

The PageRank files can be requested directly from the domain www.google.com. Basically, the URLs for those files look like follows (without line breaks):

http://www.google.com/search?client=navclient-auto&
ch=0123456789&features=Rank&q=info:http://www.do
main.com/

There is only one line of text in the PageRank files. The last cipher in this line is PageRank.

The parameters incorporated in the above shown URL are inevitable for the display of the PageRank files in a browser. The value "navclient-auto" for the parameter "client" identifies the Toolbar. Via the parameter "q" the URL is submitted. The value "Rank" for the parameter "features" determines that the PageRank files are requested. If it is omitted, Google's servers still transmit XML files. The parameter "ch" transfers a checksum for the URL to Google, whereby this checksum can only change when the Toolbar version is updated by Google.

Thus, it is necessary to install the Toolbar at least once to find out about the checksum of one's URLs. To track the communication between the Toolbar and Google, often the use of packet sniffers, local proxies an similar tools is suggested. But this is not necessarily needed, since the PageRank files are cached by the Internet Explorer. So, the checksums can simply been found out by having a look at the folder Temporary Internet Files. Knowing the checksums of your URLs, you can view the PageRank files in your browser and you do not have to accept Google's 36 years lasting cookies.

Since the PageRank files are kept in the browser cache and, thus, are clearly visible, and as long as requests are not automated, watching the PageRank files in a browser should not be a violation of Google's Terms of Service. However, you should be cautious. The Toolbar submits its own User-Agent to Google. It is:

Mozilla/4.0 (compatible; Google Toolbar 1.1.60-deleon; OS SE 4.10)

1.1.60-deleon is a Toolbar version which may of course change. OS is the operating system that you have installed. So, Google is able to identify requests by browsers, if they do not go out via a proxy and if the User-Agent is not modified accordingly.

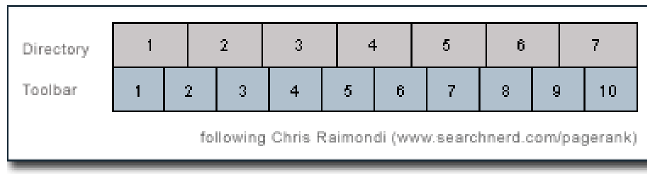
Taking a look at IE's cache, one will normally notice that the PageRank files are not requested from the domain www.google.com but from IP addresses like 216.239.33.102. Additionally, the PageRank files' URLs often contain a parameter "failedip" that is set to values like "216.239.35.102; 1111" (Its function is not absolutely clear). The IP addresses are each related to one of Google's seven data centers and the reason for the Toolbar querying IP-addresses is most likely to control the PageRank display in a better way, especially in times of the "Google Dance".

V. THE PAGERANK DISPLAY AT THE GOOGLE DIRECTORY

Webmasters who do not want to check the PageRank files that are used by the toolbar have another possibility to receive information about the PageRank of their sites by means of the Google Directory (directory.google.com).

The Google Directory is a dump of the Open Directory Project (dmoz.org), which shows the PageRank for listed documents similarly to the Google Toolbar display scaled and by means of a green bar. In contrast to the Toolbar, the scale is from 1 to 7. The exact value is not displayed, but it can be determined by the divided bar respectively the width of the single graphics in the source code of the page if one is not sure by looking at the bar.

By comparing the Toolbar PageRank of a document with its Directory PageRank, a more exact estimation of a pages PageRank can be deduced, if the page is listed with the ODP. This connection was mentioned first by Chris Raimondi (www.searchnerd.com/pagerank).



Especially for pages with a Toolbar PageRank of 5 or 6, one can appraise if the page is on the upper or the lower end of its Toolbar scale. It shall be noted that for the comparison the Toolbar PageRank of 0 was not taken into account. It can easily be verified that this is appropriate by looking at pages with a Toolbar PageRank of 3. However, it has to be considered that for verification pages of the Google Directory respectively the ODP with a Toolbar PageRank of 4 or lower have to be chosen, since otherwise no pages linked from there with a Toolbar PageRank of 3 will be found.

VI. CONCLUSION

There much more I could write about PageRank, but I hope this gives you a good introduction and some clarity about it. The key points to remember:

- PageRank tells how important a page is, relatively speaking, compared to other pages.
- PageRank is just one of MANY ranking factors used to determine ranking in search results.
- High PageRank does NOT guarantee a high search ranking for any particular term. If it did, then PR10 sites like Adobe would always show up for any search you do. They don't.
- The anchor text of a link is often far more important than whether it's on a high PageRank page.
- If you really want to know what are the most important, relevant pages to get links from, forget PageRank. Think search rank. Search for the words you'd like to rank for. See what pages come up tops in Google. Those are the most

important and relevant pages you want to seek links from. That's because Google is explicitly telling you that on the topic you searched for, these are the best.

PageRank is now one of 200 ranking factors that Google uses to determine a page's popularity. Google Panda is one of the other strategies Google now relies on to rank popularity of pages.[54] Even though PageRank is no longer directly important for SEO purposes, the existence of back-links from more popular websites continues to push a webpage higher up in search rankings.

REFRENCES

1. *The Anatomy of a Large Scale Hypertextual Web Search Engine (PDF, 1998) by Sergey Brin and Lawrence Page*

2. *What can you do with a Web in your Pocket (PS, 1998) by Sergey Brin, Rajeev Motwani, Larry Page and Terry Winograd*

3. *The PageRank Citation Ranking: Bringing Order to the Web (PDF, 1999) by Lawrence Page, Sergey Brin, Rajeev Motwani and Terry Winograd*

4. *Efficient Crawling Through URL Ordering (PDF, 1998) by Junghoo Cho, Hector Garcia-Molina and Lawrence Page*

5. *Attack Resistant Trust Metrics (PDF, 2002) Draft of PhD thesis in compact formatting by Raph Levien*

6. *Altman, Alon; Moshe Tennenholtz (2005). "Ranking Systems: The PageRank Axioms" (PDF). Proceedings of the 6th ACM conference on Electronic commerce (EC-05). Vancouver, BC. Retrieved 2008-02-05.*

7. *Sergey Brin and Lawrence Page. The anatomy of a large-scale hypertextual Web search engine. Computer Networks and ISDN Systems, 30(1–7):107–117, 1998.*

8. *Cheng, Alice; Eric J. Friedman (2006-06-11). "Manipulability of PageRank under Sybil Strategies" (PDF). Proceedings of the First Workshop on the Economics of Networked Systems (NetEcon06). Ann Arbor, Michigan. Retrieved 2008-01-22.*

9. *Langville, Amy N.; Meyer, Carl D. (2006). Google's PageRank and Beyond: The Science of Search Engine Rankings. Princeton University Press. ISBN 0-691-12202-4.*

10. *Richardson, Matthew; Domingos, Pedro (2002). "The intelligent surfer: Probabilistic combination of link and content information in PageRank" (PDF). Proceedings of Advances in Neural Information Processing Systems*

Ms. S.Sowbaranika is presently pursuing B.E Computer Science & Engineering, SNS College of Technology, affiliated to Anna University, Chennai, Tamilnadu, India. Her research interests includes Data Mining and Network Security. She has published 2 papers in National Conferences.

Mr. B.Thirunavukarasu is presently pursuing B.E Computer Science & Engineering, SNS College of Technology, affiliated to Anna University, Chennai, Tamilnadu, India. His research interests includes BigData , Data Mining, Business Analytics and Web services. He has published 2 papers in National conference and 2 papers in journals. He is an active entrepreneur involving web services and mobile application development

Mr. S. Stewart Kirubakaran is presently working as an Assistant Professor in the Department of Computer Science and Engineering, SNS College of Technology, affiliated to Anna University, Coimbatore, Tamilnadu, India. He received the M.E degree from Kalasalingam University. His research interests include network security, web services and Data Mining.

Dr.S.Karthik is presently Professor & Dean in the Department of Computer Science & Engineering, SNS College of Technology, affiliated to Anna University- Coimbatore, Tamilnadu, India. He received the M.E degree from the Anna University Chennai and Ph.D degree from Anna University of Technology, Coimbatore. His research interests include network security, web services and wireless systems. In particular, he is currently working in a research group developing new Internet security architectures and active defense systems against DDoS attacks. Dr.S.Karthik published more than 35 papers in refereed international journals and 25 papers in conferences and has been involved many international conferences as Technical Chair and tutorial presenter. He is an active member of IEEE, ISTE, IAENG, IACSIT and Indian Computer Society.

