

STUDY ON ENHANCING BLOG QUALITY USING SOCIAL CONNECTIVITY

M. POORNA CHANDAR, MAYANK SHARMA, M.V.VIJAYA SARADHI

Abstract: *Blogging has become a useful way for people to publish data on the Internet, which contains posts, comments, trackbacks etc. They can also share latest news and opinions. As several blogs have risen up on the web, the users have the problem to identify which of the blogs contain useful information & data. My attempt in this paper is to present a road map in ranking blogs. In the first blush, it appears that this exercise is the same as ranking the general web ranking but actually the blogs in Blogosphere are not amenable to using these same algorithms. An approach based on social connectivity between bloggers i.e. Relationships among bloggers and different important features could help in enhancing blog quality. In our approach we are adding "Time" Factor by analyzing these blogs relationships at different time intervals would help us to identify the impact of blogs in Blogosphere.*

Keywords: *Blogosphere, Blogs, Ranking and Analyzing.*

I. INTRODUCTION

Social networking [1] has been around long before the internet was introduced to us. Social networking is the practice of expanding the number of one's business and/or social contacts by making connections through individuals. While social networking has gone on almost as long as societies themselves have existed, the unparalleled potential of the Internet to promote such connections is only now being fully recognized and exploited, through Web-based groups established for that purpose. Online community services are sometimes considered as a social network service, though in a broader sense, social network service usually means an individual-centered service whereas online community services are group-centered.

Social networking sites allow users to share ideas, activities, events, and interests within their individual network. For example Social networking sites such as MySpace, Bebo, Twitter, Flickr, FriendWise, FriendFinder, Orkut, Facebook and etc. Jorn Barger coined the term blog in 1997. One of the Social Network examples is Blogging [2].

Manuscript received October 22, 2011.

M. POORNA CHANDAR, PG Scholar, Department of Information Technology, Aurora's Engineering College ,Bhongir, Andhra Pradesh, India.(E-mail: chandutmangipudi@gmail.com)

MAYANK SHARMA ,Associate Professor, Department of Information Technology, Aurora's Engineering College, Bhongir, Andhra Pradesh, India. (E-mail: mayank_sharma04@yahoo.com)

Dr. M.V.VIJAYA SARADHI, Professor, Department of Information Technology, Aurora's Engineering College, Bhongir, Andhra Pradesh, India.(E-mail: meduri_vsd@yahoo.co.in)

Blogging is very popular today because it allows people to interact with each others. A blog is also known as Weblog. It is a personal online diary or you can say journal which allows you to share your thoughts and ideas, you can add video, games, pictures music, you can read comments visitor leave on your blog. Each of these entries is called blog posts.

Generally blogs tend to have a few things in common:

- A main content area with articles listed chronologically, newest on top. Often, the articles are organized into categories.
- An archive of older articles.
- Any one and any number of users can leave comments about the articles.
- A list of links to other related sites, sometimes called a "blogrolls".

A post can include comments, trackbacks, citations from other bloggers, indicating user interest in that post's topic. One of the important point is to identify in the blogosphere is which blog is much more important than other blogs i.e. whether the content of the blog is useful for the user or not, because the reason that some of the algorithms are not considering content of blogs such as comments, trackbacks, blogrolls, citations, guest blogging and other different features which are important for ranking blogs. So to identify this problem we introduce a relationship between different blogs which finds the blogger quality. Such interactions or linking structure among the blogs in the blogosphere can help us to find blogger behaviour or blog quality that is individual role and impact of the blogs in the Social network.

II. PROBLEM OVERVIEW

A Blog contains posts, comments, trackbacks, blogrolls, daily diaries, description of events, and other material such as graphics or videos.

Now days it is very easy for every user to create their own blogs and can rapidly share any content on the internet such as daily diaries, discussions about the latest news and events, and even express their own opinions on different topics. So it is very difficult part for the user to identify

which blogs does contains the most useful information & data in the blogosphere because some blog ranking algorithms or blog service providers will rank blogs based on number of visitors that is for example Google. If number of hits or visitors is more for a blog than that blog will be consider as popular blog based on these types of algorithms. Although it is not a good approach for ranking blogs, because this indicates a weak and an insufficient way for identifying blog popularity, which is very important in the massive Blogosphere. The main reason is that such services or algorithms, will not consider detail information of the blog such as content, comments, trackbacks, guest blogging, citations, blogrolls and other different number of features between different Blogs.

III. RELATED WORK

Google proposed an algorithm to measure the web pages known as PageRank [4] algorithm. Google PageRank is one of the methods Google uses to determine a page's relevance or importance. Important pages receive a higher PageRank i.e. a page that is linked to by many pages with high PageRank receives a high rank itself and are more likely to appear at the top of the search results. In essence, Google interprets a link from page A to page B as a vote, by page A, for page B. If there are no links to a web page there is no support for that page. Next Jon Kleinberg a professor proposed an algorithm that made use of the link structure of the web in order to find and rank pages relevant for a particular topic which is known as Hyperlink-Induced Topic Search [5] algorithm. This algorithm models communities as interconnection between 'authorities' and 'hubs' these two scores for each page: its authority, which estimates the value of the content of the page, and its hub value, which estimates the value of its links to other pages.

Next Ko Fujimura and his colleagues were proposed a new link-analysis EigenRumor algorithm [6] for ranking blogs. In the EigenRumor model they consider only agent-to-object links, not page-to-page links, that is agent is used to represent an aspect of human being such as a blogger, and an object is used to represent any object such as a blog entity. Later Nardi et al. [7] found that many people creating blogging at the urging of friends and continue to blog to avoid disappointing their readers. Much work must still be done in examining this flourishing phenomenon as it grows and changes. Next Alvin Chin and Mark Chignell [8] have studied on how to identify communities can be discovered based on interconnections between blogs. They mainly focus for identifying community in blogs, and in similar forms of social hypertext. This method includes the use of visualizations and centrality measures based on networking structure, and measures of strength of community based on survey data. In examining the social structures created by blogrolls and blog comments, it is unclear how these relationships and social networks were formed.

Later so many questions have been raised that Why do people maintain weblogs? Do people who receive responses and interact with others remain active for longer than those

who do not? Or is it more personal, where some individuals are more internally motivated to continue posting content than others? For answering to these questions Thomas Lento [9] and his colleagues applied the logic regression model and network visualizations to analyze blog data in wallop system. The analysis focuses on the relationship between social ties and continued activity as expressed through various features of the system, and tests if people who remain active are more socially connected to other users within the system. Next Noor Ali-Hasan and Lada Adamic [10] analyzes online relationships and real life social networks. Next Tadanobu Furukawa [11] and his colleagues defined how the user's behavior in the blogosphere which is called as regular reading relation among bloggers. Their research work done on that reading is affective as writing and almost same as listening as talking. Later Eytan Adar [12] and his colleagues proposed the concept of implicit links representing similarity among blog posts and also iRank algorithm. Whereas traditional ranking strategies rely primarily on explicit link structure, iRank successfully folds in implicit routes of transmission to find blogs that are at the source of information.

Later on Apostolos Kritikopoulos [13] and his colleagues present a method for ranking weblogs based on the link graph and on several similarity characteristics between weblogs. They have proposed a method for using link graph characteristics and common attributes between the blog posts to enhance the efficiency of the ranking method for each blog importance. Later Nitin Agarwal [15] and his team members address a novel problem of identifying influential bloggers at a blog site by presenting a preliminary model of identifying influential bloggers of a community blog site. They mainly focus on Influential bloggers can exist at many blog sites, regardless of these sites being influential or not. They did so many experiments on essential issues of identifying Influential bloggers; evaluate the effects of various collectable statistics from a blog site on determining blog-post influence. Above all works focuses on the community of blogs but it is very important to identify the bloggers behavior in the Blogosphere. The blogosphere arose from the idea that all blogs exists as a heavily inter-connected community, so much so that it has become like a social network in itself.

The blogosphere [14] consists of everything that has to do with blogs: posts, comments, and its users. The existing service does not consider the relationship and interactions between different bloggers, which could really help in identifying the blog quality or its popularity. First of we need to know why some blogs are more important than others.

Our Basic idea is to create interactions between different blogs such as comments, trackbacks, blogrolls, citations, and guest blogs and also we have so many different important features which will help us to enhance the quality of a blog in the Blogosphere.

IV. ENHANCING BLOG QUALITY

In our proposed idea analyzers mainly concentrate on blog characteristics. So analyzers can assume famous and highly popular blogs will contain most useful data and informative content for readers. Posts in such blogs could include good discussions or innovative opinions, depending on subjective judgments. So in order to enhance the quality of a blog we made some assumptions in the Blogosphere:

- Popular or Highly quality blog will have more relationships with other blogs.
- Generally users might cite a blog or a blog post in other WebPages, for advertising about blog or blog posts.
- We can also assume the quality of a blogs, if it frequently referred on the web.
- So many important features which are much important for a blog are guest blogging, discussion forums, keywords and etc.

Let us discuss about all those features which will help to enhance the blog quality:

A. *Comments*: A comment is most interactive feature in Blogosphere. Most blogs have a method to allow comments. When blog A comments on blog B post blog B receive a note from blog A.

B. *Trackbacks*: It is a method of Person A saying to Person B, this is something you may be interested. To do that, the Person sends a Trackback to Person B.

- Person A writes something on their blog.
- Person B wants to comment on Person A's blog, but wants her own readers to see what she had to say, and be able to comment on her own blog
- Person B posts on her own blog and sends a trackback to Person A's blog
- Person A's blog receives the trackback, and displays it as a comment to the original post. This comment contains a link to Person B's post.

C. *Blogrolls*: It is method that where blogs provides a list of favourite blogs on the side of your blog.

D. *Pingback*: A pingback is a notification letting a blog or website knows that it has been referenced by other blog. Pingbacks usually contains a short excerpt of the post containing the link, along with a link to the website.

- Alice posts to her blog. The post she's made includes a link to a post on Bob's blog.
- Alice's blogging system contacts Bob's blogging system and says, "look, Alice made a post which linked to one of your posts!"

- Bob's blogging system then, when people view the posts on Bob's blog, notes on the page that Alice linked to this post.
- Users can then follow this link back to Alice's post and read more.

E. *Citations*: It is a relationship from blog A to blog B if an entry of blog A contains the link to blog B. These are very important in the Blogosphere.

F. *Guest Blogging*: It is a process between two bloggers for increasing the traffic to their own blogs. Guest blogging simply means the exchange of content from one blogger to another. It involves those people who post their articles to other websites other than their own. These people are known as guest bloggers.

G. *Similarity Relationships*: When we encounter a common hyperlinks between in both blog A and blog B then we create similarity relationships between these two blogs.

H. *Discussion Forums*: Blogging community showing your interest in contributing guest posts on other blogs in your niche.

I. *Permalinks*: Permalinks are the permanent URLs to your individual weblog posts. A permalink is what another weblogger will use to refer to your article, or how you might send a link to your story in an e-mail message.

These explicit and implicit relationships and interactions between different bloggers will help us to identify highly popular blog. So these interactive blog behaviors are used to construct a Social relationship network in the Blogosphere. In order to analyze the link structure of our sample we represent the blog link structure as a graph, associating several attributes with the vertices and edges. Each vertex represents data about a blog, and each edge represents data about a link including data about the post where the link was extracted from.

In this social blog network model, we have two activities such as Node and Edges:

- Node: Each node is represented as a single blog.
- Edges: Each edge between two nodes represented as a relationship between two blogs.

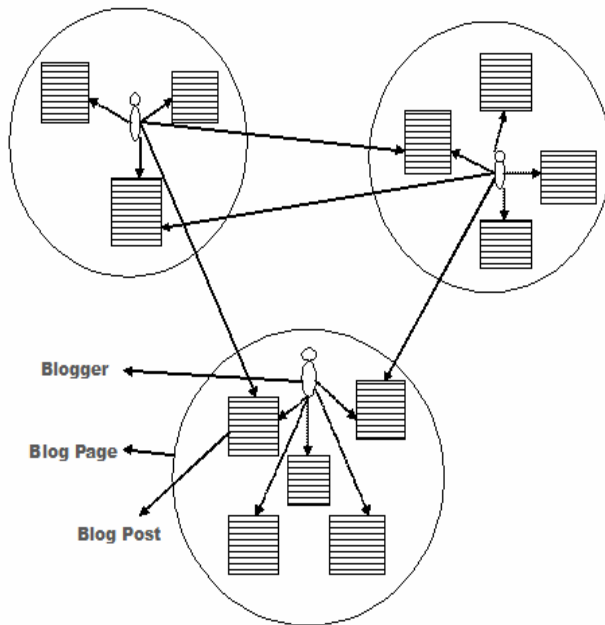


Figure 1: linking relationships between blogs.

In the above figure we represent each circle as a blog (node) and each link represents the relationships between different blogs. Interactions or links from one blog to another blog is represents that blog A is reader of blog B.

The above edges divided into various ways:

- *Support Relationships:* The Support relationships are Comments, Trackbacks, Citations, Blogrolls, and Permalinks and Guest Blogging
- *Similarity Relationships:* Similarity Relationships form similarity edges when two blogs have common interest links.
- *Hyperlink edges:* This relationship exists between blog and website or vice versa

The number of internet users is continuously expanding, and the number of people with personal blog is rising in a vertiginous rate from one year to the next. We have so many algorithms which concentrates on number of hits of the page but not content of the page such PageRank or HITS algorithm .So this is not an efficient and these indicates week for identifying Popular or Highly quality blogs in the massive Blogosphere. Our Study on various papers and blogs on the websites we have noticed that interactions or interconnections between bloggers will reveals Blogger behaviour in the Blogosphere .Not only the interactions between the blogs we also various important parameters such as keywords, RSS, Social Bookmarking and etc can also be implemented for Ranking Blogs.

There are so many important features are mentioned below:

1. *Recommendation:* How much your blog recommending other article?
2. *Keywords:* Every time you write a post you should consider what words people might be putting into search engines to find that type of information. Keeping such keywords in title is Stronger.

3. *Advertisement:* Putting some dollars on advertising today might mean that your blog gets popular a within less time. For example: Pay per click, Google adsense etc.
4. *RSS:* RSS is used to publish recently updated information in the websites. It is very important which will increase your blog’s traffic. You should submit your RSS feed to feeding websites.
5. *Famous person:* This represent is that the blogger is the Famous Person or not.
6. *Social Bookmarking:* This is another to get your website or Blog popularity. This means providing your blog to social media websites which increases your traffic and popularity might also becomes high.
7. *Frequency:* Is your blog content is frequently updating or not because the search engines tend to spider the Web pages at frequent time snapshots.
8. *Unique visitor:* A unique visitor is access from a single IP to a web server that generates page views and hits during a particular visit.
9. *Comment spam filter:* Commenting is an essential part of your blog because it creates community by allowing reader participation. However, those of you that have a blog know just how annoying blog spam can be. While many blogs use.
10. *Most popular post list:* People visiting your blog for the first time need to be directed to your most valuable posts. Providing a single area on one of your side bars with just the titles of your most trafficked posts can not only increase visitor retention but also increase subscriptions since it shows the true value of your blog.
11. *Press Release Submission:* Press Release Submission is another great way to build quality backlinks to your blog because they are providing the content from lots of minor sites. And they easily chosen by news sites that mean more and effective way to gain quality backlinks.

We also study about one parameter which is very good measurement for identifying a blogger behaviour or blog quality that is main factor “TIME” .So by analyzing our social network at different time snapshots we can observe Structural workflow of each blog .We actually assume some assumption for analyzing blog at different time frequency that means how the user is going to the web site or blog most frequently .So here we are going to find the time interval at different iterations so for that we can easily identify the most frequently access blog in the blogosphere this plays very crucial role in Social Networking.

V. CONCLUSION

We have so many Blog ranking algorithms and ranking service providers for ranking Blogs but they very week and inefficient in ranking Blogs because those service providers



or ranking websites don't consider the detail content or information such comments, trackbacks, citations and don't consider some important features which could help us to identify blogger quality or enhance blog quality in the Blogosphere. Based on many authors we conclude that interactions between different bloggers will help the analyzer to identify the blog popularity. And we study that different parameters such social bookmarking, guest blogging, RSS etc can also increase the Blog traffic easily. We also are adding one factor that is "TIME" which is very important feature which could identify blogger behavior in the Blogosphere. By analyzing the Blog at different time snapshots we can find how the user is frequently accessing the website or Blog in the Blogosphere by calculating the time interval at different iterations.

VI. REFERENCES

- [1] DiMicco, J., Millen, D., Geyer, W., Dugan, C., Brownholtz, B. 2008. Motivations for Social Networking at Work. ACM CSCW'08.
- [2] Blood, R. *Weblogs: A History and Perspective*. Rebecca's Pocket, Sept. 7, 2000; www.rebeccablood.net/essays/weblog_history.html.
- [3] Herring, S. C., Kouper, I., Scheidt, L. A., & Wright, E. (2004). Women and children last: The discursive construction of weblogs. In L. Gurak et al. (Eds.), *Into the Blogosphere: Rhetoric, Community, and Culture of Weblogs*. http://blog.lib.umn.edu/blogosphere/women_and_children.html
- [4] L. Page et al., *The PageRank Citation Ranking: Bringing Order to the Web*, tech.report, Stanford, InfoLab, 1999; <http://ilpubs.stanford.edu:8090/422/1/1999-66.pdf>
- [5] J.M. Kleinberg, "Authoritative Sources in a Hyperlinked Environment," *J. ACM*, vol. 46, no. 6, 1999, pp. 604-632.
- [6] K. Fujimura, T. Inoue, and M. Sugisaki, "The EigenRumor Algorithm for Ranking Blogs," *Trusting Agents for Trusting Electronic Societies*, LNCS 3577, Springer, 2005, pp. 59-74.
- [7] B.A. Nardi et al., "Why We Blog," *Comm. ACM*, vol. 47, no. 12, 2004, pp. 41-46.
- [8] A. Chin and M. Chignell, "A Social Hypertext Model for Finding Community in Blogs," *Proc. 17th Conf. Hypertext and Hypermedia*, ACM Press, 2006, pp. 11-22.
- [9] T. Lento et al., "The Ties that Blog: Examining the Relationship Between Social Ties and Continued Participation in the Wallop Weblogging System," *Proc. Workshop on the Weblogging Ecosystem: Aggregation, Analysis and Dynamics*, ACM Press, 2006; www.blogpulse.com/www2006-workshop/papers/Lento-Welser-Gu-Smith-TiesThatBlog.pdf.
- [10] N. Ali-Hasan and L.A. Adamic, "Expressing Social Relationships on the Blog through Links and Comments," *Proc. Int'l Conf. Weblogs and Social Media (ICWSM 07)*, AAAI Press, 2007; www.icwsm.org/papers/2--Ali-Hasan--Adamic.pdf.
- [11] T. Furukawa et al., "Social Networks and Reading Behavior in Blogosphere," *Proc. Int'l Conf. Weblogs and Social Media (ICWSM)*, AAAI Press, 2007; www.icwsm.org/papers/2--Furukawa-Matsuo-Ohmukai-Uchiyama-Ishizuka.pdf.
- [12] E. Adar et al., "Implicit Structure and the Dynamics of Blogspace," *Proc. Workshop on the Weblogging Ecosystem: Aggregation, Analysis and Dynamics*, ACM Press, 2004; www.blogpulse.com/papers/www2004adar.pdf.
- [13] A. Kritikopoulos, M. Sideri, and I. Varlamis, "BlogRank: Ranking Weblogs Based on Connectivity and Similarity Features," *Proc. 2nd Int'l Workshop Advanced Architectures and Algorithms for Internet Delivery and Applications*, ACM Press, no. 8, 2006; <http://doi.acm.org/10.1145/1190183.1190193>.
- [14] Chih-Lu Lin and Hung-Yu Kao, "Blog Popularity Mining Using Social Interconnection Analysis" Published by the IEEE Computer Society, 41-49, 2009; http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=5432133.
- [15] N. Agarwal et al., "Identifying the Influential Bloggers in a Community," *Proc. Int'l Conf. Web Search and Web Data Mining*, ACM Press, 2008, pp. 207-218.
- [16] M.A. Tayebi, M.S. Hashemi, and A. Mohades, "B2Rank: An Algorithm for Ranking Blogs Based on Behavioral Features," *Proc Int'l Conf. Web Intelligence*, IEEE CS Press, 2007; <http://doi.ieeecomputersociety.org/10.1109/WI.2007.81>.



M. POORNA CHANDAR is currently pursuing M.Tech from Aurora's Engineering College, Bhongir, Andhra Pradesh, India. Contact him at chandutmangipudi@gmail.com



MAYANK SHARMA has been working as an Associate Professor at Aurora's Engineering College, Bhongir, Andhra Pradesh, India. Contact him at mayank_sharma04@yahoo.com



Dr. M.V.VIJAYA SARADHI has been working as a Professor at Aurora's Engineering College, Bhongir, and Andhra Pradesh, India. Contact him at meduri_vsd@yahoo.co.in