

Searching and Rating of Trusted Blogs on Web Pages

S.Aruna Devi

Assistant Professor

*Department of Computer Science and Engineering
Saveetha Engineering College, Chennai, Tamilnadu, India*

D.Ramalingam

Assistant Professor

*Department of Computer Science and Engineering
Saveetha Engineering College, Chennai, Tamilnadu, India*

Abstract- A weblog or blog in short is a special web page on which a individual author (blogger) or a group of periodically publish articles (entry/post). Web blogs are prominent social media on the internet which allows bloggers and users to interact efficiently with the knowledge, personal details etc., Because of the popularity and easy access of blogs most of the people started using blogs for various purposes. Since the blogs are open and free to all anyone can write and read a blog. The problem is that how to trust the blog published by multi group of bloggers. There are huge numbers of blogs scattered in the web on various topics. The objective of our proposed system is to create trust among bloggers and users by analyzing explicit and implicit characteristics of a blogger. The distributed blogs are collected from the web using crawler and create semantic annotation among the blogs. The blogs are classified according to topics and personalized according to user profile and general profile of the user. The blogs can also be ranked by considering user search, corpus quality, feedbacks, popularity etc., thus the system provide efficient scalable semantic linked personalized blogs to the user.

Keywords – Trust, Blog, Personalization, Semantic annotation.

I. INTRODUCTION

The blog is a new internet tool that gives users a channel to express their knowledge and feelings with other people worldwide. A blog is a website consisting of posts generated dynamically and updated frequently, presented in reverse chronological order and written by a single person or group. Blogs enable Internet users to either express their ideas or to share knowledge and thoughts with other users worldwide. Blogging software enables Internet users to publish their works without technical or editorial constraints. Blogs count increase rapidly at the rate of 1.4 new blog per second. Blogosphere is the social network formed by the interactions between authors and readers. More blogs are not well-connected properly or completely unconnected and a blogger may write on variety of subjects[2]. Hence, the readers have difficulty in finding blogs with articles that might fit their tastes or information needs based on the links between blogs in the blogosphere. Since bloggers cover diverse subjects, and because many blog readers are unskilled in using the Internet, bloggers and readers experience difficulty in finding valuable articles from the hundreds of millions of blogs of the Internet.

In order to address the challenges imposed by recent explosion of blogs, search tools that can help users efficiently access and discover blog posts become highly necessary. Searching blog posts by use of existing popular web search engines however may produce unsatisfactory results, since blogs possess several unique characteristics that differentiate them from general web pages. It often contains a blogroll, which is a list of blogs favorite blogs that the blogger reads regularly. The activity involving in maintaining and updating the blog are called blogging, and the blog owner is called blogger (blog author). Bloggers and blog readers usually interact via their blogs and through related functions such as commentaries and blog rolls. Trust is a belief and willingness to act on the basis of the words, actions, and decisions of others, based on positive expectations of their intentions or behaviors. Blogs are increasingly accepted as a useful means to proliferate a variety of information on the web.

II. PROPOSED SYSTEM

Personalization Of Blogs: The proposed system aims to exploit the trust relationships between bloggers and readers via explicit trust ratings to generate recommendations in a reliable and satisfactory way. It presents a multi-faceted model that considers trust by dividing a general trust rating into multiple trust ratings for different types of blog articles, considering the number of blogs posted by the blogger, comments of various users etc., thus enabling trust relationships to be evaluated in a fine-grained manner so that the blogs can be effectively retrieved based on the trust value. It creates the semantic relationship between blogs so that it overcomes the difficulty in finding out blogs that fit the users' needs since the blogs which are relatively likely in its contents get integrated. It creates a significantly positive correlation between trust and similarity in blogosphere. The blogs are then retrieved based on the general and personal profile of the blogger.

The main contributions of this paper are to enable i) integration of all relevant topics.

ii) retrieval of trusted blogs in efficient and easier way iii) retrieval of blogs based on general and personal profile of the user thereby making the user query to be satisfied in a better way.

III. EXPERIMENT AND RESULT

The overview of the paper is shown in Figure 1. The system has six modules viz. user profiling, blog creation, retrieval of trusted blogs, semantic annotation between blogs, ranking of blogs and personalization of blogs. A database is adopted to store the data given by the user, system data, blog articles, comments of user etc.

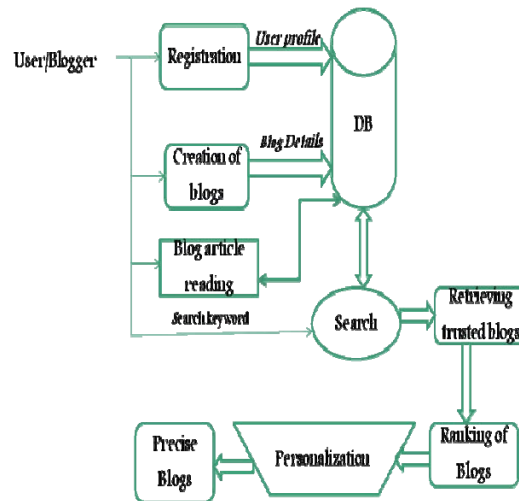


Fig. 1 Architecture of the system

Initially a user must register for his account by using the user profiling module that contains a list of general and personal information. This module validates the user with his inputs and stores the data provided by the user into database after successful validation. Blog creation module performs the functionalities such as creating blogs by the blogger, viewing blogs both by the blogger and user, commenting the blogs, adding new blog category to existing blog name. The created blogs are stored into database so that when the searching process proceeds the blogger can read the blog article and can provide comments. Retrieval of trusted blogs depend on the trust rating of the user which gets calculated by considering the number of blogs posted by the blogger, comments for his blogs, rating given to its contents. Since the blogger while registering gets validated, these parameters enable to verify whether the blog is trusted and when the blog article gets searched the blogs are retrieved based on the trust value. Semantic annotation of blog module is used to create relationship among relevant blogs thereby enable the availability of strongly connected blogs. The module, ranking is used to order the blogs based on its content similarity to the search query given by the user. Finally, Personalization of blogs makes the user satisfied with what information he needs based on his profile thereby filtering out irrelevant data provided to the user.

Only users with levels of trust greater than a threshold in particular types are considered as comparable neighbors for generating recommendations. The detailed algorithm is described as follows. If D is defined as the set of

categories in blog article taxonomy and C_k as a category in D , then t_{x,y,C_k} denotes the explicit trust rating in category C_k rated on a scale of 1 to 10.

$$t_{x,y,C_k} \in [+1, +10], C_k \in D$$

By considering the diverse of trust ratings by users, trust ratings need to be normalized to a specific range. To do this Max-Min normalization is applied to derive $trust(x,y,C_k)$, which indicates the level of trust between users x and user y in category C_k and has values on a scale of 0-1.

$$trust(x,y,C_k) = \frac{t_{x,y,C_k} - \min_x^t}{\max_x^t - \min_x^t}$$

Significantly, since trust and similarity exhibit a [positive correlation this study only considers users with both positive similarity ($sim(x,y) \geq 0$) and sufficient trustworthiness ($trust(x,y,C_k) > T$) to be the comparable neighbors for the computation of recommendation. Therefore, $Y^T(i)$ denotes the most similar and trustworthy users profiles with trust values $trust(x,y,C_k)$ of category C_k exceeding a predefined threshold T .

$$Y^T(i) = \{y \in Y(i) : trust(x,y,C_k) > T, sim(x,y) \geq 0\}$$

IV.CONCLUSION

This paper paid attention to the blogs in the blogosphere. It solves the problem of availability of incredible blogs thereby enabling the improvement in user satisfaction by providing the blogs according to its trust value and it also performs the functionality of personalization in which it provides only the needed information to the user's query. Comments to which trusted blocks got with high rating posted by users will be displayed on top of the web pages. This is going to be a future work.

REFERENCES

- [1] P.Avesani, P.Massa and R.Tiella. A trust-enhanced recommender system application: Moleskiing. *Proceedings of the 2005 ACM symposium on Applied computing*, 2005. 1589-1593.
- [2] M.Dell'Amico and L.Capra. SOFIA: Social Filtering for Robust Recommendations. *International Federation for Information Processing (IFIP), Trust Management II*, 263, 2008. 135-150.
- [3] S.Godfrey Winster, S.Swamynathan, "Blog Trust Model for Blog Readers", *Proceedings of IEEE International Conference on Recent Trends in Information, Telecommunication and Computing*, pp 314-317, 2010.
- [4] S.C.Herring, I.Kouper, J.C.Paolillo, L.A.Scheidt, M.Tyworth, P.Welsch, E.Wright and N.Yu. Conversations in the Blogosphere. An Analysis from the BottomUp. *Proceedings of the 38th Hawaii International Conference on System Sciences (HICSS'05)*, 2005.
- [5] N.Lathia, S.Hailes and L.Capra. Trust-based Collaborative Filtering. *International Federation for Information Processing (IFIP), Trust Management II*, 263, 119-134, 2008
- [6] P.Massa and P. Avesani. Trust aware recommender systems. *Proceedings of the 2007 ACM conference on Recommender systems*, 17-24, 2007
- [7] Nanno, T. Fujiki, T. Suzuki, Y., and Okumura, M. .Automatically collecting, monitoring and mining Japanese Weblogs. *13th International conference*, 320-321, 2004
- [8] P.Resnick and H.R.Varian. Recommender systems. *Communications of the ACM*, 40, 33, 1997. 56-68
- [9] Seung- Kyun Han, Dongmin Shin. Jae-Yoon Jung, Jonghun Park, "Exploring the relationship between keywords and feed elements in Blog post search", 2009
- [10] Ting-Chung Peng, Seng-cho T.Chou "iTrustUA Blog Recommender System based on Multi-faceted trust and Collaborative Filtering", pg no. 1278-1285, 2010
- [11] Bloglines: Available: <http://www.bloglines.com>
- [12] www.flash.com
- [13] Yun chi, Belle L.Tseng, Junichi Tatemura, "Eigen-Trend Trend analysis in the Blogosphere based on singular value decomposition", *International conference on Information and Knowledge Management CIKM'06*, pp 391-397, 2006
- [14] C.N. Ziegler and J.Golbeck. "Investigating interactions of trust and interest similarity. *Decision Support Systems*", 43 ,22, 460-475, 2007.