Survey of Investigation on Ranking of Active Web Forum Users

Shreethika S S[1], Reshma P[2], Aishwarya M[3], Indira Kumari B[4]
[1,2,3,4] UG Scholar, Department of IT, Karpagam Institute of Technology, Coimbatore, Tamilnadu, India.

Abstract: *The World Wide Web consists of millions of interconnected web pages that provide information to the user present in any part of the world. The World Wide Web consists of the web forums, where active Web users participate and post their comments. The influential users dominate and influence the newly joined innocent users through their thoughts. The activeness of a user is captured by a measure based on the degree of match of the commented posts with a complete list. Hence it is necessary to retrieve the best and the influential user who frequently participate and post their right message that are more relevant in terms of information for the query entered by the user in the search engine. The proposed algorithm is weighted ranking algorithm for accurate identification of active users. This includes about five metrics to identify the influential users. The process starts with a data crawling and preprocessing step in which the URL of the forum home page is passed to the forum crawler, which crawls all relevant web pages and eliminates the duplicates heuristically. Users can use the web services and provide their feedbacks on the forum. All these data's are stored in the log file. Their comments are compared with the threat list. From the list, the most influential users are listed out based on the session count and the session time of users.*

Keywords: *User analysis, security informatics, user collocation analysis, active user identification.*

INTRODUCTION

In the recent past it is found that the world wide web consists of many interconnected web pages that provide information to the user in any part of the world. The world wide web consists of many web forums where people can participate and post their comments for the particular event or the product. The infiltration of unwanted comments and the action which is posted by the influential user to attract and dominate the naive user is considered as a major threat or noise to the user who view those comments. Web forums are recognized for their exhaustive, vivid and non-spontaneous nature of discussions that are archived for later reference .Previous studies have found Web forums as the most active medium being used for this purpose. Research on identifying active and infectious members, postings, ideas and ideologies in Web forums for tracking the grievous threats posed by the active extremist and hate groups has gained considerable attention of the research community. It is necessary to remove the unwanted comments in the forum and order the best comments through comparison of the created table and rank the users who frequently participate in posting the proper comments for the benefits of the user. Those users are called active users who give right comments to the naive users. All those users should be authorized in order to rank the active users based on session count.

Existing System

With the rapid growth of user-generated contents, the study of information propagation and influential users in particular web forum has become crucial to a plethora of related analysis problems. This section presents some of the important previous works on influential user identification.

Cite this article as: Shreethika S S, Reshma P, Aishwarya M, Indira Kumari B. "Survey of Investigation on Ranking of Active Web Forum Users". *International Conference on Systems, Science, Control, Communication, Engineering and Technology 2016*: 498-499. Print.

A majority of previously studied works on the problem of influential user identification have been done in a business intelligence orientation for marketing products through targeted influential users or viral marketing [8], [12]. Some other objectives are information dissemination [13], community leader identification [14], and expertise discovery [15].Richardson and Domingos [12] worked on the social network formed from collaborative ratings, and modeled it as Markov random fields, considering each customer's product buying probability as a function of both its intrinsic desirability for the customer and the influence of others. Ghosh and Lerman [16] utilized the dynamics of voting on *digg* posts to rank influential users. They defined an empirical measure of influence based on the number of in-network votes that the post of a user receives. Kempe *et al.* [17] devised a greedy approach based discrete-optimization model to maximize the spread of influence through a social network. However, Kimura *et al.* [13] found that the computational cost of a conventional greedy approach to identify influential nodes in a network is very high, and consequently they proposed a method of estimating marginal gains on the basis of bond percolation and graph theory. Hill *et al.* [18] performed a statistical analysis on *email* network-based marketing and established a hypothesis for a direct affect of network linkages on product/service adoption. Java *et al.* [19] applied the influence models proposed by [17], in addition to applying algorithms like PageRank, in blogosphere. They also discussed the importance of splog removal and its implications on influence models. Agarwal *et al.* [9] came up with a comprehensive definition of influential bloggers and the challenges associated with their identification. Using an influence graph of blog posts, they defined some measures to find influential blog-posts and bolggers. Zhang *et al.* [15] proposed ExpertiseRank to rank the Java expertise using forum threads and posts in the popular Java Forum. Tang and Yang [20] contributed towards online healthcare social networks, specifically the Swine Flu online forum which is a sub-community of MedHelp. Based on the concepts of PageRank algorithm, they proposed User Rank to identify the influential users using content similarity and response immediacy. It is shown as out-performing PageRank, in-degree and out-degree rankings. In [11], they also showed the application of UserRank algorithm in the domain of Web forums.

References

1. J. Qin, Y. Zhou, and H. Chen, "A multi-region empirical study on the Internet presence of global extremist organizations," *Inf. Sys. Frontiers*, vol. 13, no. 1, pp. 75–88, 2011.
2. T. Anwar and M. Abulaish, "Modeling a Web forum *ecosystem* into an enriched social graph," in *Ubiquitous Social Media Analysis*. Berlin, Germany: Springer-Verlag, 2013, pp. 152–172.
3. T. Anwar and M. Abulaish, "Identifying cliques in Dark Web forums—an agglomerative clustering approach," in *Proc. IEEE ISI*, Jun. 2012, pp. 171–173.
4. H. Chen, W. Chung, J. Qin, E. Reid, M. Sageman, and G. Weimann, "Uncovering the Dark Web: A case study of Jihad on the Web," *J. Amer. Soc. Inf. Sci. Technol.*, vol. 59, no. 8, pp. 1347–1359, 2008.
5. J.-H. Wang, T. Fu, H.-M. Lin, and H. Chen, "A framework for exploring Gray Web forums: Analysis of forum-based communities in Taiwan," in *Proc. IEEE Int. Conf. ISI*, May 2006, pp. 498–503.
6. J. Qin, Y. Zhou, E. Reid, G. Lai, and H. Chen, "Analyzing terror campaigns on the Internet: Technical sophistication, content richness, and Web interactivity," *Int. J. Human-Comput. Stud.*, vol. 65, no. 1, pp. 71–84, 2007.
7. J. Glaser, J. Dixit, and D. P. Green, "Studying hate crime with the Internet: What makes racists advocate racial violence?" *J. Soc. Issues*, vol. 58, no. 1, pp. 177–193, 2002.
8. M. Trusov, A. V. Bodapati, and R. E. Bucklin, "Determining influential users in Internet social networks," *J. Marketing Res.*, vol. 47, no. 4, pp. 643–658, Aug. 2010.
9. N. Agarwal, H. Liu, L. Tang, and P. S. Yu, "Identifying the influential bloggers in a community," in *Proc. Int. Conf. WSDM*, 2008, pp. 207–218.
10. D. J. Watts, "Challenging the influential hypothesis," *WOMMA Meas. Word Mouth*, vol. 3, pp. 201–211, autumn 2007.
11. C. C. Yang, X. Tang, and B. M. Thuraisingham, "An analysis of user influence ranking algorithms on Dark Web forums," in *Proc. ISI-KDD*, 2010, Art. ID 10.
12. M. Richardson and P. Domingos, "Mining knowledge-sharing sites for viral marketing," in *Proc. 8th ACM SIGKDD Int. Conf. KDD*, 2002, pp. 61–70.
13. M. Kimura, K. Saito, and R. Nakano, "Extracting influential nodes for information diffusion on a social network," in *Proc. 22nd Nat. Conf. AAAI*, 2007, pp. 1371–1376.
14. I. Esslimani, A. Brun, and A. Boyer, "Detecting leaders in behavioural networks," in *Proc. Int. Conf. Adv. Soc. Netw. Anal. Mining (ASONAM)*, Aug. 2010, pp. 281–285.
15. J. Zhang, M. S. Ackerman, and L. Adamic, "Expertise networks in online communities: Structure and algorithms," in *Proc. 16th Int. Conf. WWW*, 2007, pp. 221–230.