

# Ranking Radically Influential Web Forum Users on Social Media

Samiksha R. Tadas<sup>1</sup>, Prof. K. B. Bijwe<sup>2</sup>

<sup>1</sup> P. G. Student, P. R. Pote(Patil) Welfare & Education Trust's college of Engineering & Management, Amravati  
Department of Computer Science & Engineering,

<sup>2</sup>Assistant Professor, P. R. Pote(Patil) Welfare & Education Trust's college of Engineering & Management, Amravati  
Department of Computer Science & Engineering

\*\*\*

**Abstract** - Social media is the term used to describe the interaction between groups and individual in which they produce share and sometime exchange over idea over internet and in virtual communication. The impact of social networking is on young people and children. Online social networking is spread day-by-day for various online community purposes. Now a day's social media has been the important part of one's life from shopping to electronic mails, education and business tool. Social media plays a vital role in transforming people's life style. Social media includes social networking sites and blogs where people can easily connect with each other. Since the emergence of these social networking sites like Twitter and Facebook as key tools for news, journalists and their organizations have performed a high-wire act. These sites have become a day to day routine for the people. Some of the web forums are predominantly being used for open discussions on critical issues influenced by radical thoughts. The influential users dominate and influence the newly joined innocent users through their radical thoughts. Some forum are used for open discussions on a critical topics influenced by radical thoughts. The influential users dominate the mind of naive users using their radical thoughts. Influential users complete the naive users to do wrong things. The main intension of this topic is to understand blocking of malicious links on social media post using content mining and text comparison with existing item sets which will predicate the category of the post and help to stop violent data on social media.

**Key Words:** Social media analysis, security informatics, radical user identification, user collocation analysis.

## 1. INTRODUCTION

Social media has integrated itself into many aspects of our lives. Whether we use it to bridge friendships, stay informed, or express ourselves, social media offers us a unique way to stay connected with the world around us. Now days, Web is being used as a tool to practice several kinds of mischievous acts with concealed agendas and promote ideologies in a sophisticated manner. Infiltration of extremist groups, hate

groups, racial supremacy groups, and terrorist organizations on the Web with hundreds of multimedia websites, online chat rooms and Web forums is posing grievous threats to our societies as well as the national security.



Fig 1: Online Social Media

The multimedia websites provide support for their psychological warfare, fund-raising, recruitment, and propagation of their agendas whereas chat rooms and Web forums promote their strategies and ideologies through discussions with naive users. Often the public discussions among differently minded extremist groups lead to irascible talks accompanied with abusive languages, and promote online hate and violence. Web forums are recognized for their exhaustive, vivid and non-spontaneous nature of discussions that are achieved for later reference. Previous studies have found Web forums as the most active medium being used for this purpose. Research on identifying radical and infectious threads, members, postings, ideas and ideologies in Web forms for tracking the grievous threats posed by the active extremist and hate groups has gained considerable attention of the research community. The portion of the Web circumscribing the sinister objectives of extremists group is said as Dark Web, and specifically the Web forums with substantial prevalence of activities supporting extremism are said as Dark Web forums. Another class called Gray Web forums refer to the forums in which the discussions focus on topics that might potentially

encourage biased, offensive, or disruptive behaviors and may disturb the society or threaten public safety. They include topics like pirated CDs, gambling, spiritualism, bullying, and online-pedophilia.

Due to enormous and rapid growth of user-generated content on social media sites, and users generally avoid going through every comment posted by others. There always exist some users who develop some relationship of trust with other members by their activeness and quality of comments, and their comments always receive significant attention of a large community. These are the influential users, whose activities and comments greatly affect the society. Influential users find it very easy to convince the other users with their ideologies. Recent studies have found it to be an important issue and a challenging task to identify such influential users. Influential user identification has been done in a business intelligence orientation for marketing products through targeted influential users. Some other objectives are information dissemination community leader identification and expertise discovery. An empirical measure is done of influence based on the number of in-network votes that the post of a user receives content similarity and response immediacy. It is shown as out-performing PageRank, in-degree and out-degree rankings helps in the identification of the user and the application of User Rank algorithm in the domain of Dark Web forums.

## 2. EXSITING SYSTEM

Social media is a form of internet marketing that utilizes social networking websites as a marketing tool. The goal of social networking is to produce content that users will share with their social network help a company increase brand exposure and broaden customer reach. In the recent past, it has been found that the web is also being used as a tool by radical or extremist groups and users to practice several kinds of mischievous acts with concealed agendas and promote ideologies in a sophisticated manner. Some of the web forums are predominantly being used for open discussions on critical issues influenced by radical thoughts. The influential users dominate and influence the newly joined innocent users through their radical thoughts. The radical of a user is captured by a measure based on the degree of match of the commented posts with a threat list. Eleven different collocation metrics are formulated to identify the association among users, and they are finally embedded in a customized Page Rank algorithm to generate a ranked list of radically influential users. Social media is very popular for different kind of activities. In this system web forum users contain for complete analysis of forums for manipulating the information post by users in forum. User of such forum talk about the violence data through their post, there are numbers of website are also available that are spreading such data. The reality is that system not successfully defined violence present

on the social media. Existing system works only crawling the forums but proposed system crawling normal user post and helps to stop violence data on social media. So that violent posts do not attack on people. The proposed system also finds out average percentage value of violent data present on the social and developed an effective for healthy society through healthy social media environment.

## 3. PROPOSED SYSTEM

Social media is the process of marketing through social media sites like twitter, Facebook and you tube. By utilizing the social aspect of the web, social media marketing is able to connect and interact on a much more personalized and dynamic level than through traditional marketing. Online Social Networks is mainly used as an interactive medium to communicate, share, a considerable amount of human life information. It is a platform to build social networks or social relations among people, for example share interest, pictures, and text and real time connections. Social media is term used to describe the interaction between groups and individual in which they produce share and something exchanged idea over the internet and in virtual communication. The impact of social media on the young people significant. Children are growing around surrounded by mobile device and interactive social networking sites such as twitter, MySpace & Facebook. So that to avoid this drawback and improve violence posting on the social media. The proposed system which will check for the user post links post by them and their content in post as well as link and determining the abusing of that which help to make social media to make more secured from spreading such type of volatile thoughts on web.



Fig 3.1: Social Networking

Social networking sites render the opportunity for people to reconnect with their old friends, colleagues and mates. It also helps people to make new friends, share content, pictures, audios, videos amongst them. Social media also changes the life style of a society. One of the negative effects of social media is that it makes people addicted. People spend lots of

time in social networking sites which can divert the concentration and focus from the particular task. Social media can easily affect the kids.

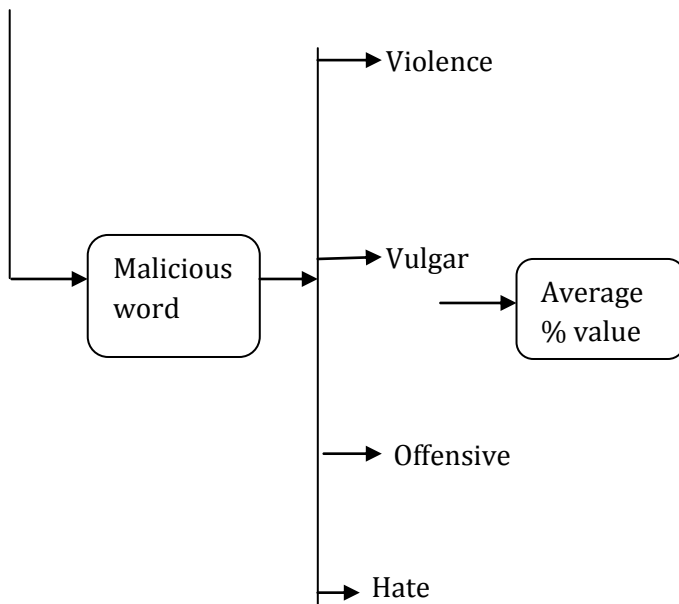


Fig 3.2: Diagram of working system

In the above figure, first the user is login and set the post. The posts contain only textual data. Splitter is used to spread out word to word data from the user post and then malicious word detected. There are four types of categories of the post such as violence, vulgar, offensive and hate. In this system which will check for the user post such as texts or links. Determining the category of the post and find out the average percentage value of malicious post on the social media which make more secured from spreading such type of violence thought on the social networking. Social media also changes the life style of a society Another positive effect of social networking sites is it unite people on a huge platform for the achievement of specific goals. This brings positive change in the society. Social media provides awareness among society like campaigns, advertisement articles, promotions which helps the society to be up to date with the current information.

#### 4. IMPLEMENTATION

The best entertainment for the younger generation is given in the form of Social Networking sites. The Online Social Networks mainly helps an individual to connect with their friends, family and the society online in order to gather and share new experiences with others. Now days, the Online Social Networks are facing the problem of the people posting the indecent messages on any individual's wall which annoys other people on seeing them. The proposed system performs blocking the malicious data on the social networking sites.

1. Firstly in proposed method first it fill-up the registration form
2. Secondly login the system.
3. Now in third step user set the post. The post contain only textual data.
4. In Fourth step system perform average percentage value of the violent data and also find out the categories of the post i.e., violence, vulgar, offensive or hate.
5. After that in fifth step admin wall shows the result which are view user, view all posts, user malicious post calculator.
6. In the step all results are displayed which are the output of above all step.
7. This is the last step in that apply finally blocking the malicious post in the social networking and helps to stop the violent data. The system make more secured for spreading such type of volatile thought on wed.

#### 5. RESULT ANALYSIS

##### A] Inserting post or links on wall of social networking site

When user's inserting or posting the statement that time system providing the check post, the check post will identifying that inserted post is statement or link. After identifying things the system forwarding control to next level. In this section it is also important to understand the type of posting, there is several kind of posting we can do on the wall of social networking sites.

##### B] Extraction of words

After identifying whether it is link or statement of post, here systems will applying the data mining techniques. System will be using data mining classifications techniques to extraction of words from the statement of the post, and web mining algorithm system using for web links to collect the information's of the web link, and after collecting all information's here system will checking out that particular web links or statements is malicious or not.

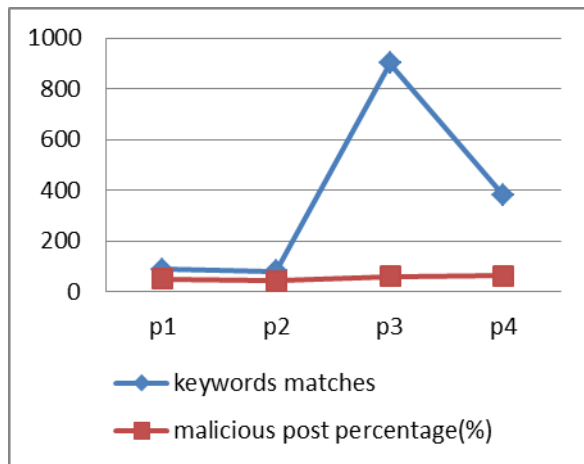
##### C] Detections of Malicious Post

After applying the data mining techniques and web mining algorithm, here system checking the maliciousness of the post with existing blacklisted keywords which is present in the database, if the calculated post is malicious and more than 67.33% malicious then it will be directly blacklisted, else post maliciousness is less than 67.33% then it will be available on wall of the users. To calculating the percentage of the malicious link or post we are using the following formula.

$$\text{Average \% value} = \frac{f(T)+f(L)}{2} * 100$$

Where, f(T)= Number of item set element in text

$f(L)$ = Number of item set element in link



**Graph 5.1:** User post malicious percentage

In above analysis of graph and considered user posts P1, P2, P3, P4. For every user of post we are calculating the malicious percentage with the help of keywords matches, here we get the results that are for user P1 88 keywords are matched with the database blacklisted keywords and after applying the malicious formula we got results which is shows in graph. The graph clearly shows that post of the user is malicious the red line in graph shows malicious percentage of post, and blue line shows that keywords matches with the database blacklisted keywords.

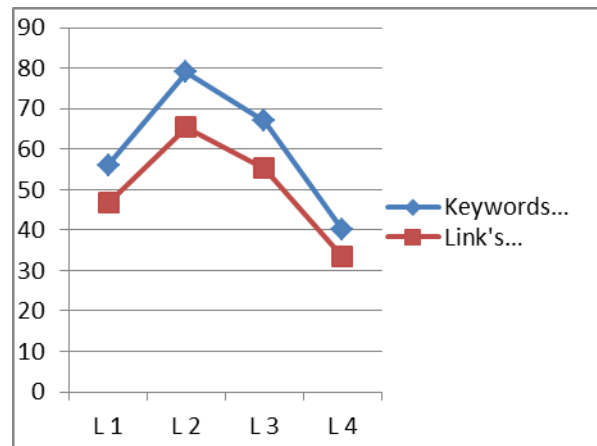
User link	Keyword matches	Malicious
L1	56	46.66
L2	79	65.33
L3	67	55.33
L4	40	33.33

**Table 5.2:** User’s link malicious percentage analysis

In above table showing analysis of graph 7.2 and considered user posts L1, L2, L3, L4. For every user of post we are calculating the malicious percentage with the help of keywords matches, here we get the results that is for user L1 56 keywords are matched with the database blacklisted keywords and after applying the malicious formula we got results which is shows in above table.

We consider for study blacklisted keywords in database is 120, on that basis we are calculating the malicious percentage of the link, here the blacklisted keywords in database is continuously updating so for study we are considering a fix value for database blacklisted keywords that is 120. Out of 120 here for every users of malicious link will matches the pattern of keywords with database blacklisted keywords. Depending on the keywords matches

the graph showing the results. No. of blacklisted keywords are matches more than malicious percentage of link also increase, so we can say that malicious percentage is directly depending on the blacklisted keywords of the database.

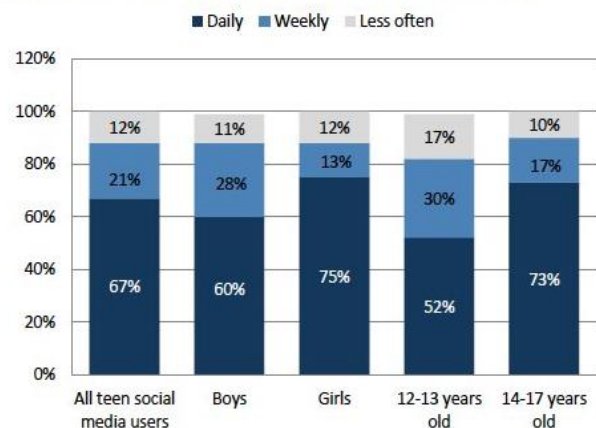


**Graph 5.3:** User’s link malicious percentage analysis

The graph clearly shows that post of the user’s is malicious, the red line in graph shows malicious percentage of post, and blue line shows no. of keywords matches with the database blacklisted keywords.

**Daily use of social media**

% of teen social media users who visit social networking sites, by frequency



**Fig 5.4:** Daily use of social media

### Younger social media users are more likely to use the tools for civic activities

% of users of social networking sites or Twitter who use social media these ways

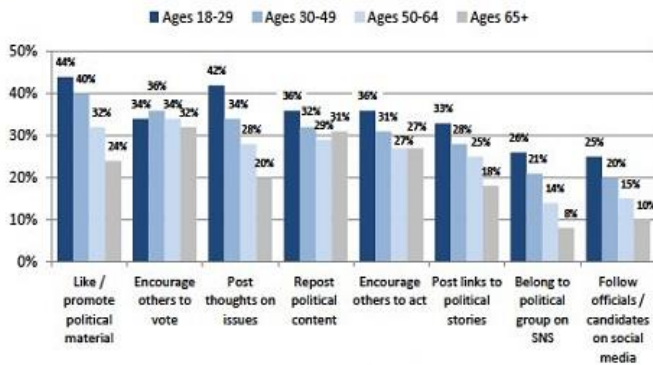


Fig 5.5: Social Networking Activities

Now a day's most of the applications are running on internet as online social networking. The proposed work can be implemented over the social networking to stop the violence post. By implementing such work which help to make more secured from the violence data present on the social media. People have formed online communities that allowed them to get support, education and even promote and sell products.

### 6. CONCLUSION

As the technology is growing the social media has become the routine for each and every person, people are seen addicted with these technology every day. With different fields its impact is different on people. Social media has increased the quality and rate of collaboration for students. Business uses social media to enhance an organization's performance in various ways such as to accomplish business objectives, increasing annual sales of the organization. Young people are seen in contact with these media daily .Social media has various merits but it also has some demerits which affect people negatively. False information can lead the education system to failure, in an organization wrong advertisement will affect the productivity. This technique helps to stop violence data on the social media. Here literature study is performed on the different techniques used by different researchers in the field and then proposed system Determining the category of the post and find out the average percent value of malicious post on the social media. So that social media helps young people to stay connected with each other. Useful information can be exchanged over social networking sites. Young people can look to social media for getting the answers related to their career objectives.

### 7. FUTURE SCOPE

Social media is term used to describe the interaction between group and individual in which produce share and sometime exchange idea over the internet and in virtual communication. Children are growing around surrounded by mobile device.so that the proposed system stops violent data on social media. Social media can provide a supportive environment to explore friendship, and social status, while also providing an opportunity to share and discuss their knowledge. Blocking the violence post on social media provides a clear platform for educating employees.

### ACKNOWLEDGEMENT

The author would like to present their sincere gratitude towards the, Prof. K. B. Bijwe (Co Guide), Prof. P. P. Pawade (Guide) and also thanks to Prof. V. B. Gadicha (H.O.D – Computer Science & Engineering) for their extreme support to complete this assignment.

### REFERENCE

- [1] Samiksha Tadas and Prof. K. B. Bijwe "A Survey on Ranking Radically Influential Web Forum Users on Social Media", International Journal of Innovative Science, Engineering & Technology, vol. 3, Issue 11, November 2016
- [2] Tarique Anwar and Muhammad Abulaish, "Ranking Radically Influential Web Forum Users", IEEE Transaction on information forensics and security, Vol. 10, No. 6, June 2015
- [3] Priyanka B. Mane, Sonali S. Rathid, Deepali D. Sanap & Bhavana S.Shirude, "Ranking Influential Users", Imperial Journal of Interdisciplinary Research (IJIR), Vol-2, Issue-5, 2016
- [4] M. Richardson and P. Domingos, "Mining knowledge-sharing sites for viral marketing", In Proc. 8th ACM SIGKDD Int. Conf. KDD, pp. 61-70 ,2012.
- [5] 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.
- [6] Esslimani, A. Brun, and A. Boyer, "Detecting leaders in behavioral networks", In Proc. Int. Conf. Adv. Soc. Newt. Anal. Mining (ASONAM), pp. 281-285, Aug. 2010.

- [7] 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.
- [8] Denzil Correa, Ashish Sureka, "Solutions to Detect and Analyze Online Radicalization", In *Proc. IIITD PhD Comprehensive Report*, Vol. V, No. N, November 2011
- [9] T. Anwar and M. Abulaish, "Identifying cliques in Dark Web forums—An agglomerative clustering approach", In *Proc. IEEE ISI*, pp. 171–173, June 2012.
- [10] 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
- [11] X. Tang and C. C. Yang, "Identifying influential users in an online healthcare social network", in *Proc. IEEE Int. Conf. ISI*, pp no43–48, May 2010.
- [12] 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.
- [13] T. Anwar and M. Abulaish, "Modeling a Web forum ecosystem into an enriched social graph," Berlin, Germany: Springer-Verlag, pp. 152–172, 2013.
- [14] T. Anwar and M. Abulaish, "Identifying cliques in Dark Web forums An agglomerative clustering approach", in *Proc. IEEE ISI*, pp. 171–173, June 2012.
- [15] 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.
- [16] 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.
- [16] 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.
- [17] Shabnoor Siddiqui, Tajinder Singh, "Social Media its Impact with Positive and Negative Aspects", *International Journal of Computer Applications Technology and Research*, Volume 5– Issue 2, 71 - 75, 2016, ISSN:- 2319–8656.