Impact Factor (SJIF): 5.301

International Journal of Advance Research in Engineering, Science & Technology

e-ISSN: 2393-9877, p-ISSN: 2394-2444
Volume 6, Issue 12, December-2019

BOTNET DETECTION BY USING MACHINE LEARNING TECHNIQUE

Sandhya Balu Dahiwade[1], Surekha Bhairu Gaikwad [2]
Prof. Nilesh Korade [3]

sandhyadahiwade123@gmail.com[1], surekha.gaikwad.pccoer@gmail.com[2],
nilesh.korade@pccoer.in[3],

[1]Computer Engineering, Pimpri Chinchwad College of Engineering And Research, Ravet, Pune University
[2]Computer Engineering, Pimpri Chinchwad College of Engineering And Research Ravet, Pune University
[3]Computer Engineering, Pimpri Chinchwad College of Engineering And Research, Ravet, Pune University

***Abstract* — A social network sites plays very important role for spreading the information and influence in the form of whole world in hand. There is a basic thing to find small set of influential people in a social network such that targeting them initially. It will increase the spread of the influence but the problem is that to finding the most influential nodes in network. There is one Algorithm called as Greedy algorithm used for mining top-K influential nodes. It consists of 2 modules: which bifurcate the post of social network into several variants by taking into account information diffusion and selecting communities to find influential nodes by a advance programming. Using Location Based Community Greedy algorithm is used to find the influence node based on Location of the user and we can sort out the influence for particular Area. There is one more Serious problem in social network is that there are so many malicious attacks spread by the people and that malicious contents are hidden behind some attractive posts. i.e viral marketing techniques in the promotion of new products or posts like how much luckiest today or what does your name meant or post which states the giving you iPhone in lesser prize or giving them free samples of the product etc. such posts in social network may be harmful to users which has the negative intention of stealing the personal information user. So the takes the advantage of such usage history of ads i.e. rating of such posts the system analyses the that products impact on social media users and predicts the positive or negative category for that posts which is beneficial for future users on social media. So we need to find out that post and block them.**

.

***Keywords: Social network, Influence maximization, Greedy algorithm, Malicious Attack*.**

I. INTRODUCTION

Today’s Social Networks are becoming fast and dynamic platforms for sharing post as well as it have become globalized market for advertisement of products. Social Media used by millions of user all over the world. But there are some posts in social network may be harmful to users which have the negative intention of stealing the personal information user. So we have to protect the users form such type of Harmful Post as well as we need to find out Owner of this post. Viral marketing techniques in the promotion of new products or posts like how much luckiest today or what does your name meant or post which states the giving you ipad in lesser prize or giving them free samples of the product etc. such posts in social network may be harmful to users which has the negative intention of stealing the personal information user. So the takes the advantage of such usage history of ads i.e. rating of such posts the system analyses the that products impact on social media users and predicts the positive or negative category for that posts which is beneficial for future users on social media.

In this work system is going to develop a system of efficient categorization technique for identifying whether a post generated by a third party application is malicious or not. Detecting malicious URLs is now an essential task in network security intelligence. To maintain efficiency of web security, these malicious URLs have to be detected, identified as well as their corresponding links should be found out. Hence due to this users get alerted and protected from it and effectiveness of network security gets increased. The malicious users can upload a content he wants to spread i.e.

malicious posts on social media. The content that contains malicious data is posted to other user's wall under a different form. The user mistakes the posts for a real content and clicks the post, which will take him to another page. Thus the malicious user can benefit from this process. In order to get the attention of the user, the malicious user will include keywords or description of pages that will be of interest to the user. These can be adult content or free downloading sites. So it needs a technique which banned such an activity by just identifying and blocking that malicious content directly.

II. LITERATURE SURVEY

I. Identifying influential nodes based on network representation learning in complex networks

Author: Wei H, Pan Z, Hu G, Zhang L, Yang H, Li X,

Description:
Network representation learning aims at learning distributed vector representation for each vertex in a network. It is also increasingly recognized as an important aspect for network analysis. In recent years, many methods have been put forward to find influential nodes in complex networks. The knowledge of node's spreading ability shows new insight for application such as controlling propagation of messages and rumors in social networks ranking reputation of scientists and finding social leaders. In this paper, they have propose an effective method based on network representation learning. The method considers not only the overlapping communities in networks, but also the network structure. Experiments on real-world networks show that the proposed method outperforms many benchmark algorithms and can be used in large-scale networks.

2. Influential Node Tracking on Dynamic Social Network: An Interchange Greedy Approach

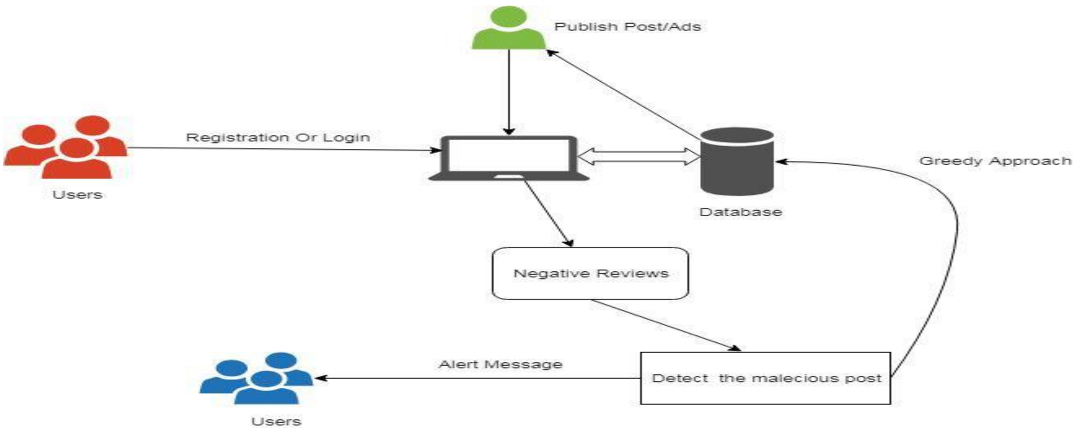
Author: Guojie Song, Yuanhao Li , Xiaodong Chen, Xinran He, Jie Tang

Description:
In this paper they have explore a novel problem, namely Influential Node Tracking problem, as an extension of Influence Maximization problem to dynamic networks, which aims at tracking a set of influential nodes dynamically such that the influence spread is maximized at any moment. It propose UBI+ algorithm that improves the computation of the upper bound and achieves better influence spread.

III. PROPOSED SYSTEM

This proposed system aims at providing highly efficient system that detects the malicious posts on social media. As a part of our system we will be developing a Social Media Site like a Facebook. Where user will be able to post large number of posts. If there is post which creates violence or post which is harmful (Malicious) the user can find out the owner of the post for that we have used Greedy approach algorithm. After that we have provided option for reviews for all posts where user can put reviews for that post and using that reviews our system will identify the post is negative or positive. If the particular post is having large number of negative reviews then our system will show popup message as alert message to the user.

I. SYSTEM ARCHITECTURE



MATHEMATICAL MODEL

Step 1: User 'Ui' will registered to 'S'.

Step 2: User 'Ui' will see the post of ads 'Ai' on this timeline.

Step 3: User will give rating 'Ri' (rating are like 1 to 5 points).

Depending on the usage history of particular ad 'Ai' by users system will apply the efficient algorithm to detect the influence and the category of that ad.

Here, the category may be P or N which is calculated by average of particular app
being used by users.

Avg = (sum of R) / total number of that ads users.

if avg Avg is greater than threshold average value then that ad post is considered as positive category else
it is negative.

Step 4: As per Negative Rating System will Notify to new user. ie Alert about malicious Post.

Step 5: System will Block that Post.

IV. APPLICATIONS

- Find the Post Owner on Social Media.
- Identify Malicious Post by Reviews.
- Block Malicious Posts.
- Showing Alert Message for Malicious Post.

V. HARDWARE REQUIREMENT

Hard Disk	: 40 GB.
System	: Intel I3.
Monitor	: 15 VGA Colour.
Ram	: 4 GB.
Mouse	: Logitech.

VI. SOFTWARE REQUIREMENT

Operating system	: Windows XP Professional/7LINUX.
Coding language	: JAVA/J2EE.
IDE	: Eclipse Kepler.
Database	: MYSQL,XAMPP

VII.CONCLUSION

We have explained how the influence maximization problem in Social Network can be analyzed using simple data taken from Social Network. We are using the greedy approach algorithm method for finding the owner of the Post/add in a social network. This Searching Technique allows to be finding owner of the post/add which are containing malicious data/links. Therefore, the Node which are having malicious content is considered as most harmful node which influences another node in the network. When the malicious node is identified then the next step is finding communication of the node network and how many nodes are influenced by that malicious node. After that we will get reviews of that node and if particular post is very dangerous then we are blocking that post, So we can avoid malicious attacks on another system and Social Media will become Secure.

.

VIII.REFERENCES

[1] W. Chen, Y. Wang, and S. Yang, “Efficient influence maximization in social network,” in KDD, 2009, pp. 199–208.

[2] P. Domingos and M. Richardson, “Mining the network value of customers,” in KDD, 2001, pp. 57–66.

[3] D. Kempe, J. Kleinberg, and E. Tardos, “Maximizing the spread of inffluence through a social network,” in KDD, 2003, pp. 137–146.

[4] M. Kimura and K. Saito, “Tractable models for information diffusion in social networks,” in PKDD, 2006, pp. 259–271.

[5] W.Yu, G.Cong, G.Song, and K.Xie, “Community-based greedy algorithm for mining top-k influential nodes in mobile social networks,” in KDD, 2010, pp. 1039–1048.

[6] W.Chen, C.Wang, and Y.Wang, “Scalable influence maximization for prevalent viral marketing in large-scale social networks,” in KDD, 2010, pp. 1029–1038.

[7] W. Chen,W. Lu, and N. Zhang, “Time-critical influence maximization in social networks with time-delayed diffusion process,” in AAAI, 2012.

[8] W. Chen, Y. Yuan, and L. Zhang, “Scalable influence maximization in social networks under the linear threshold model,” in Data Mining (ICDM), 2010 IEEE 10th International Conference on. IEEE,2010, pp. 88–97.

[9] N. Du, L. Song, M. Gomez-Rodriguez, and H. Zha, “Scalable influence estimation in continuous-time diffusion networks,” in Advances in neural information processing systems, 2013, pp. 3147–3155.

[10] J. Leskovec, J. Kleinberg, and C. Faloutsos, “Graphs over time:densification laws, shrinking diameters and possible explanations,”in KDD, 2005, pp. 177–187.

[11] J. Leskovec, J. M. Kleinberg, and C. Faloutsos, “Graph evolution:Densification and shrinking diameters,” TKDD, vol. 1, 2007.

[12] J. Leskovec, L. Backstrom, R. Kumar, and A. Tomkins, “Microscopic evolution of social networks,” in KDD, 2008, pp. 462–470.

[13] C. Zhou, P. Zhang, J. Guo, X. Zhu, and L. Guo, “Ublf: An upper bound based approach to discover influential nodes in social networks,” in ICDM, 2013.