Influential Node Tracking on Dynamic Social Network using Data Mining

Neha Tiwari[1] Amruta Chore[2]
[1,2]Assistant Professor
[1]Department of Computer Engineering [2]Department of Electronics & Telecommunications Engineering
[1,2]Dr. D. Y. Patil Institute of Engineering Management & Research Akurdi, Pune, India

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 has two components: dividing the post of social network into several communities by taking into account information diffusion and selecting communities to find influential nodes by a dynamic 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.

Key words: 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. The advertisers have started their promotions on social media to attract new customers around the globe. 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.

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. Searching and finding malicious or harmful URLs or links is now an difficult task in network security intelligence. To maintain efficiency of web security, these malicious/harmful urls needs to be detected, identified as well as their corresponding links should be found out. Hence users will get protected from attacks and effectiveness of network security will be increased.

II. LITERATURE SURVEY

1) Title: Scalable influence maximization in social networks under the linear threshold model
– Authors: W. Chen, Y. Yuan, and L. Zhang.
– Description

Influence maximization, defined by Kempe, Kleinberg, and Tardos, is the problem of finding a small set of seed nodes in a social network that maximizes the spread of influence under certain influence cascade models. The scalability of influence maximization is a key factor for enabling prevalent viral marketing in large scale online social networks. Prior solutions, such as the greedy algorithm of Kempe et al. (2003) and its improvements are slow and not scalable, while other heuristic algorithms do not provide consistently good performance on influence spreads. In this paper, we design a new heuristic algorithm that is easily scalable to millions of nodes and edges in our experiments. Our algorithm has a simple tenable parameter for users to control the balance between the running time and the influence spread of the use the font closest in appearance to Times. Avoid using bit-mapped fonts. True Type 1 or Open Type fonts are required. Please embed all fonts, in particular symbol fonts, as well, for math, etc.

Algorithm. Our results from extensive simulations on several real-world and synthetic networks demonstrate that our algorithm is currently the best scalable solution to the influence maximization problem: (a) our algorithm scales beyond million-sized graphs where the greedy algorithm becomes infeasible, and (b) in all size ranges, our algorithm performs consistently well in influence spread it is always among the best algorithms, and in most cases it significantly outperforms all other scalable heuristics to as much as 100increase in influence spread.

2) Title: Simulated Annealing Based Influence Maximization in Social Networks
– Authors: Qingye Jiang,Guojie Song, Gao Cong, YuWang, Wenjun Si, Kunqing Xie
– Description

The problem of influence maximization, i.e., mining top-k influential nodes from a social network such that the spread of influence in the network is maximized, is NP-hard. Most of the existing algorithms for the problem are based on greedy algorithm. Although greedy algorithm can achieve a good approximation, it is computational expensive. In this paper, we propose a totally different approach based on Simulated Annealing (SA) for the influence maximization problem. This is the first SA based algorithm for the problem. Additionally, we propose two heuristic methods to accelerate the convergence process of SA, and a new method of computing influence to speed up the proposed algorithm. Experimental results on four real networks show that the proposed algorithms run faster than the state-of-the-art

greedy algorithm by 2-3 orders of magnitude while being able to improve the accuracy of greedy algorithm.

III. PROPOSED SYSTEM

In proposed work is first of all we are going to developed a Social Media Site such as Facebook. And we will post large number of posts. If there is post which creates violence or post which is harmful (Malicious) the we need to find out owner of that post So, For Finding the Owner of Post or add we are going to use Greedy Approach Algorithm. After that we will Categorized the post by positive and negative post 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. If there is Large Number of negative Comments for particular Post or add then our system will Identify that post is malicious then our system will generate alert message for that particular post and System will block that post.

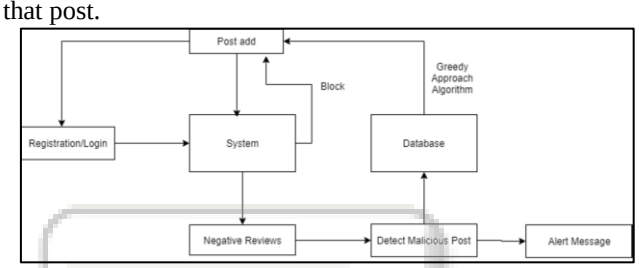

Fig. 1: System Architecture

A. *Advantages of Proposed System*

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

IV. MATHEMATICAL MODEL

Let W be the set of whole system which consists of the input, process and output of the system.
W = input, process, output.
Where,
Input = is the set of inputs given to the system to achieve the problem statement.
Process = is the procedure or the algorithm applied to the system which gives the expected output.
Output = is the output of the system.
Input = S, U, A, R, P, N, Avg.
Let,

1) S = be the social media system like Facebook.
2) U = be the set of users on social media.
   U = u1, u2, u3. .. . .un.
3) A = be the set of ads or post on social media.
   A = a1, a2, a3, ...an.
4) R = ratings given to ad A by user U
   R = r1, r2, r3, ...rn.
5) P be the positive category of ads.
6) N be the negative category of ads.
7) Avg is the average calculated by usage history particular ad A

A. *Process:*

1) Step 1: User 'Ui' will registered to 'S'.
2) Step 2: User 'Ui' will see the post of ads 'Ai' on this timeline.
3) 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.

4) Step 4: As per Negative Rating System will Notify to new user. ie Alert about malicious Post.
5) Step 5: System will Block that Post.

V. CONCLUSION

In this paper, 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.

REFERENCES

[1] W. Chen, Y. Wang, and S. Yang, Efficient influence maximization in social network, in KDD, 2009, pp. 199208.
[2] P. Domingos and M. Richardson, Mining the network value of customers, in KDD, 2001, pp. 5766.
[3] D. Kempe, J. Kleinberg, and E. Tardos, Maximizing the spread of influence through a social network, in KDD, 2003, pp. 137146.
[4] M. Kimura and K. Saito, Tractable models for information diffusion in social networks, in PKDD, 2006, pp. 259271.
[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. 10391048.
[6] W.Chen, C.Wang, and Y.Wang, Scalable influence maximization for prevalent viral marketing in large-scale social networks, in KDD, 2010, pp. 10291038.
[7] W. Chen,W. Lu, and N. Zhang, Time-critical influence maximization in social networks with time-delayed diffusion process, in AAAI, 2012.

