Recognition of Malicious Posts among Facebook Social Network Groups to Investigate User's Behavior

Sanjeev Dhawan[1], Kulwinder Singh[2], Sanjay Sagwal[3]
[1, 2] (Faculty of Computer Science and Engineering, Department of Computer Science and Engineering, University Institute of Engineering and Technology, Kurukshetra University, Kurukshetra, 136119, Haryana, India)
[3](M. Tech. [Computer Engineering], University Institute of Engineering and Technology, Kurukshetra University Kurukshetra, 136119, Haryana, India)

Abstract – Online Social Networks (OSNs) are widely used by people in today's environment. People can connect with their friends and share text, images and videos with them, also they may add their own new friends, also they may join existing group or pages according to their interest. In this paper an attempt has been made to propose a framework to distinguish genuine posts or malicious posts shared or posted by users on a Facebook page in Facebook and similarly it may be extended to other social networking sites for example(Twitter, Instagram, Whatsapp etc). To analyze proposed framework real dataset is collected from netvizz a faceb0ook application and process it using Gephi tool.

Index Terms – Online Social Networks (OSNs), Facebook, Posts, Netvizz and Gephi.

1. INTRODUCTION

In recent years, online social networks have grown rapidly and today offer individual's endless possibilities for publicly expressing themselves, communicating with friends, and sharing information with people across the world [1]. As online social networks (OSNs) are becoming the new epicenter of the web, hackers are expanding their territory to these services. Anyone using Facebook is likely to be familiar with what we call here socware1: fake, annoying, possibly damaging posts from friends of the potential victim [2]. The propagation of socware takes the form of postings and communications between friends on OSNs. Users are enticed into visiting suspicious websites or installing apps with the lure of false rewards (e.g., free iPads in memory and they unwittingly send the post to their friends, thus enabling a viral spreading. This is exactly where the power of socware lies: posts come with the implicit endorsement of the sending friend [3].

The Facebook terminology

Facebook is the largest online social network today with over 900 million registered users, roughly half of whom visit the site daily. Here, we discuss some standard Facebook terminology relevant to proposed work [4] [5].

i. Post: A post represents the basic unit of information shared on Facebook. Typical posts either contain only text (status updates), a URL with an associated text description, or a photo/album shared by a user.

ii. Wall: A Facebook user's wall is a page where friends of the user can post messages to the user. Such messages are called wall posts.

iii. News feed: A Facebook user's news feed page is a summary of the social activity of the user's friends on Facebook. For example, a user's news feed contains posts that one of the user's friends may have shared with all of their friends.

iv. Like: Like is a Facebook widget that is associated with an object such as a post, a page, or an app. If a user clicks the Like widget associated with an object, the object will appear in the news feed of the user's friends and thus allow information about the object to spread across Facebook.

v. Comments: it shows how many people interest of users on a post either positively or negatively. User can write its own feelings about that post on user's wall.

vi. Shares: It shows how many users share a post on their facebook wall.

vii. Reactions: reaction is a new widget added by facebook. A user can react on post by clicked on reaction and express its views that either smiling or crying etc.

viii. Application: Facebook allows third-party developers to create their own applications that Facebook users can add.

2. RELATED WORK

To steal user's information, attackers are creating so many fake posts and these fake posts seem to be look like real posts. That is the main aspect, many researchers and organizations are designing different techniques to protect the user from the

attackers and spammers. Therefore, Puttaswamy explained that the attacks of social intersection were an efficient and less costly to get private information of the user. Rahman *et al*. [9] developed FRAppE, a suite of efficient classification techniques for identifying whether an app is malicious or not. Smith *et al.* [10] defined Life Logging as "the collection of data in order to illustrate a person's life." In other Social Networks, such as, e.g., Twitter or Google+, Graph *G* can be modeled as a directed graph as the user connections are not necessarily bidirectional.

3. PROPOSED WORK

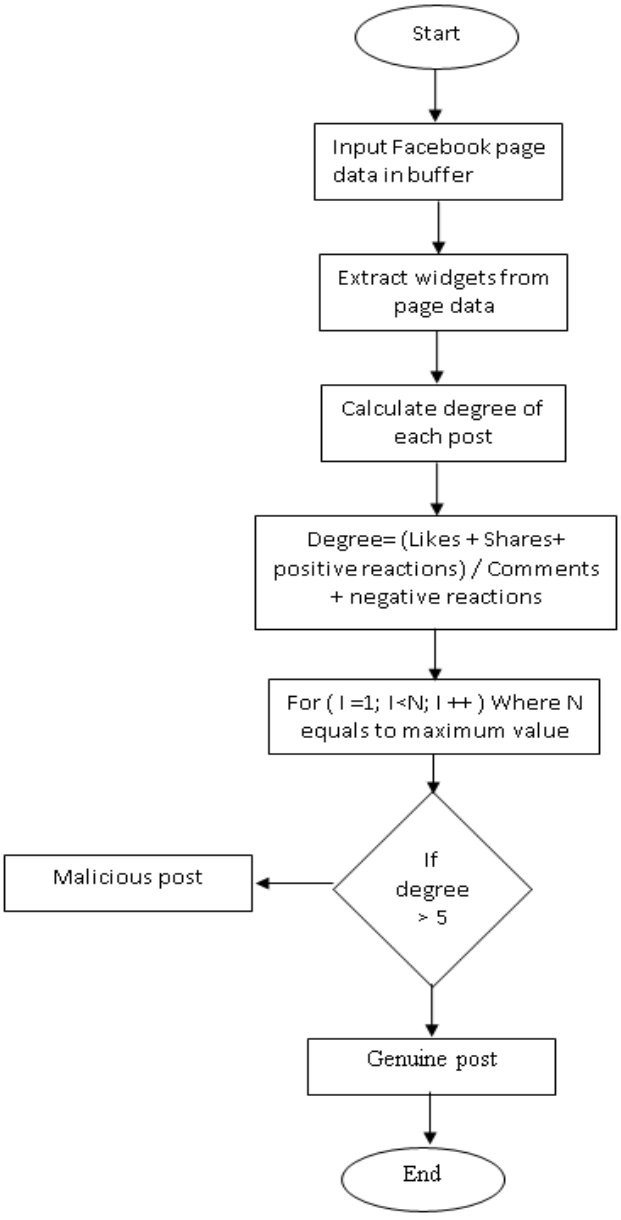

Fig. 1 Proposed Framework for Identification of Malicious Posts.

Earlier Rahman *et al.* [6] proposed malicious content detection technique. It identified malicious content and applied SVM classifier to find malicious content. Prateek *et al.* [7] proposed a technique to detect malware on Facebook using real time identification modules. Both techniques have some drawbacks such as first technique just uses SVM classifier to classify malicious or normal content. Another technique detects only malware users based on their profile information.

In Facebook Social Network user can post number of posts including text messages, images and videos and friend of users or friend of friends can like posts or they may react on that post also if they didn't like that post they may comment on it and react negatively [11, 12, 13]. So in this paper propose a framework to find which post are genuine posts or which fake or unwanted posts. To distinguish them degree of each post is calculated. The degree of post is calculated by the use of facebook widgets likes, shares, comments, positive reactions and negative reactions. If post degree is greater than some threshold value then it is said to be a genuine post otherwise it is fake or unwanted post.

$$Degree = \frac{\text{Likes} + \text{shares} + \text{Positive Reviews}}{\text{Comments} + \text{Negative Reviews}}$$

Here in this flow chart it describes that in the start of the flow chart, facebook page data from the buffer is inputted and widgets from the page data is extracted. After that degree of each post is calculated. Parameters used to calculate degree are like, comment, share and reaction. In this way degree of all the post are calculated. After that if degree of each post is greater than 5 then post is malicious otherwise the post is genuine.

4. RESULTS AND ANALYSIS

Gephi is used to visualize social network in the form of nodes and edges. In proposed work, nodes are the posts and edges are the interaction between posts. It is open source software and supports java. In gephi visualize classes used in social network without writing code [8].

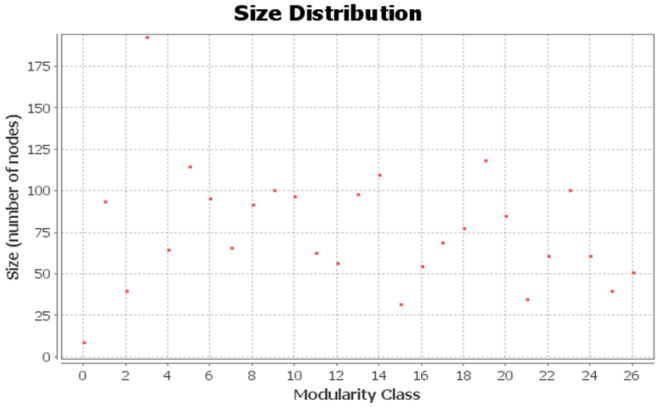

Figure 2 Modularity Classes

Fig 2 shows moduarity classses. Modualrity is used to find and distinuigh the nodes which are closely connected with each other. This size distribution is calculated on the basis of size (number of nodes) and the modularity class.

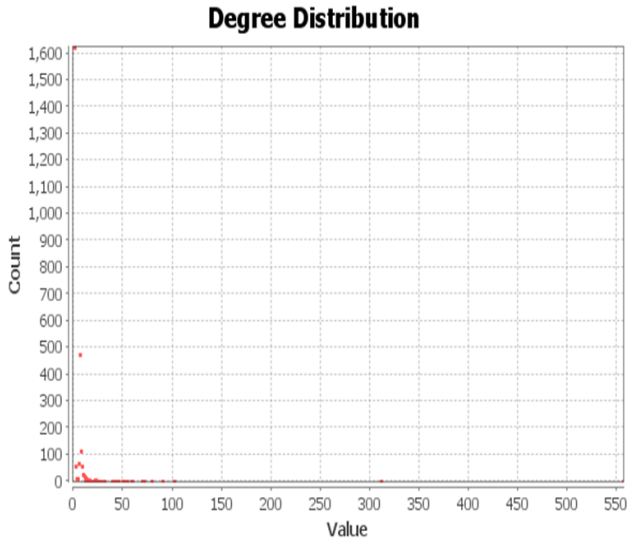

Figure 3 Degree Distribution

Fig 3 shows degree distribution of nodes. Degree distribution is used to measure how much nodes are connected with edges and by counting total number of edges to which node is connected degree of node is calculated.

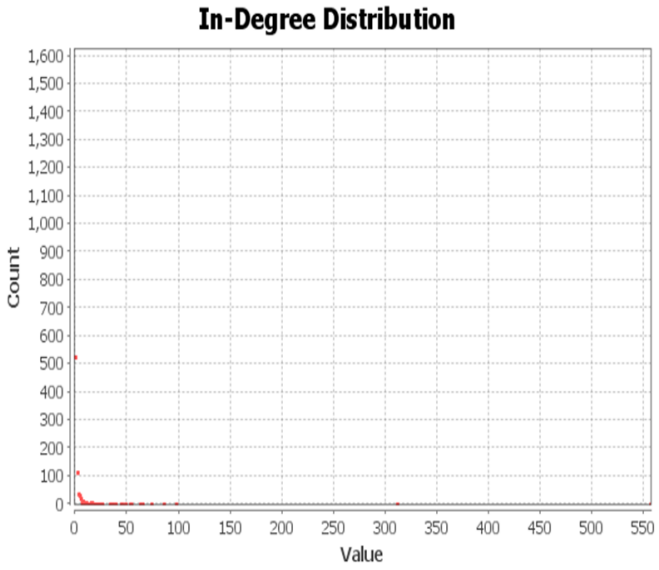

Figure 4 In-Degree Distribution

Fig 4 shows total number of incoming connections to a node. In-degree defines the no. of incoming edges to a node. In this graph in-degree is calculated with the help of count and value.

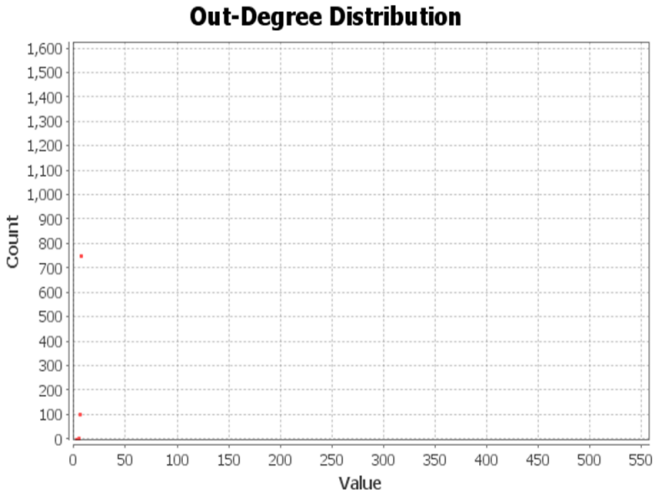

Figure 5 Out Degree Distribution

Fig 5 shows total number of outgoing connections from a node. Out-degree defines the no. of outgoing edges to a node. Out-degree is calculated with the help of count and value.

Figure 6 Recognition of Genuine Posts or Malicious Posts

Fig 6 shows recognition of genuine posts or malicious posts. In this fig maroon color show genuine posts and Faroese color shows malicious posts. So it is clear that in proposed work genuine posts are high. If genuine posts are high then it means these have more likes and shares as well as more positive reactions.

5. CONCLUSION

In Facebook social networks, as number of users increases the number of irrelevant posts are also increases. It is difficult to find which post is genuine post or not posted by facebook page admin or users who are the members of that page. So to handle this kind of problem, in this paper a framework is proposed in which based on different attributes like reactions, shares, posts and likes of a post a degree is calculated and based on that degree a post is distinguished whether it is genuine post or malicious post. Results shows that proposed framework distinguish malicious posts in more accurate manner.

REFERENCES

[1] F. Celli, "Unsupervised Personality Recognition for Social Network Sites," in ICDS 2012, The Sixth International Conference on Digital Society, no. c, 2012, pp. 59–62.

[2] Y. Bachrach, M. Kosinski, T. Graepel, P. Kohli, and D. Stillwell, "Personality and Patterns of Facebook Usage," in proceedings of the 3rd annual ACM web science conference, 2012, pp. 24–32.

[3] J. Staiano, F. Pianesi, B. Lepri, and A. Pentland, "Friends don t Lie - Inferring Personality Traits from Social Network Structure," in Proceedings of the 2012 ACM conference on ubiquitous computing, 2012, pp. 321–330.

[4] Pran Dev, Jyoti, Dr. Kulvinder Singh and Dr. Sanjeev Dhawan, "A Naive Algorithmic Approach for Detection of Users' with Unusual Behavior in online Social Networks" International Journal of Software and Web Sciences (IJSWS), ISSN: 2279-0071pp: 37-41,2015.

[5] Ekta, Sanjeev Dhawan and Kulvinder Singh, "Feature Extraction and Content Investigation of Facebook User's using Netvizz and Gephi", Advances in Computer Science and Information Technology (ACSIT), ACSIT 2016, pp. 262-265.

[6] M. S. Rahman, T.-K. Huang, H. V. Madhyastha, and M. Faloutsos, "Efficient and scalable socware detection in online social networks". In USENIX Security Symposium, pages 663–678, 2012.

[7] Prateek Dewan and Ponnurangam Kumaraguru. "Towards Automatic Real Time Identification of Malicious Posts on Facebook", 2015 IEEE, pp. 85-92.

[8] https://gephi.org/users/ accessed on 5 Aug 2017.

[9] Sazzadur Rahman, Ting-Kai Huang, Harsha V. Madhyastha, and Michalis Faloutsos, "Detecting Malicious Facebook Applications", IEEE/ACM TRANSACTIONS ON NETWORKING, IEEE 2015, pp. 1-15

[10] Smith A, O'Hara K and Lewis P, "Visualizing the past: Annotating a life with linked open data", in: Web Science Conference '11, 2011.

[11] Joe, M. Milton, and B. Ramakrishnan. "Enhancing security module to prevent data hacking in online social networks." Journal of Emerging Technologies in Web Intelligence 6.2 (2014): 184-191.

[12] Joe, M. Milton, B. Ramakrishnan, and R. S. Shaji. "Prevention of losing user account by enhancing security module: A facebook case." journal of emerging technologies in web intelligence 5.3 (2013): 247-256.

[13] Joe, M. Milton, and B. Ramakrishnan. "Novel authentication procedures for preventing unauthorized access in social networks." Peer-to-Peer Networking and Applications 10.4 (2017): 833-843.