

AN EFFICIENT TECHNIQUE FOR DETECTION OF MALICIOUS NODES IN SOCIAL NETWORKS

Ms. Hemasri Atturu¹

3rd Year Student,

Department of Computer Science,

SV U CM & CS, Tirupati.

Prof. Sridevi²,

Professor,

Department of Computer Science,

SV U CM & CS,, Tirupati.

Abstract: The anomaly detection is the technique which is applied to detect malicious activities from the social network data. The existing technique is based on to classify the Facebook accounts into three classes which are fake, genuine and moderate. To increase accuracy of account classification is increased when bloom filter is being applied in the algorithm. The bloom filter is the algorithm which learns from the previous experiences and drive new values. When the bloom filter is applied the accounts are classified into two classes. The simulation is being performed in MATLAB and it is being analyzed that accuracy is increased and execution time is reduced.

General Terms Anomaly detection, Bloom Filter, Classification, Fake accounts, Online social Networks.

Keywords Anomaly, Analysis, Classification.

INTRODUCTION

Social network analysis (SNA) is the process of investigating social structures using network and graph theories [1]. It portrays networked structures in terms of nodes (individual actors, people, or things within the network) and the ties, edges, or links (relationships or interactions) that connect them. Divisions of actors into clubs or "sub-groups" can be an essential part of social structure. It can be imperative in understanding how the network all in all is probably going to act. The idea of a clique is relatively simple. And no more general level, a clique is a sub-set of a network in which the actors are all the more closely and intensely tied to each other than they are to different members of the network. The strongest possible definition of a clique is some number of actors (more than two, usually three is used) who have every single possible tie present among themselves [2]. A Maximal Complete Sub-Graph is such a grouping, expanded to incorporate as numerous actors as possible. The N-clique

approach tends to discover long and stringy groupings instead of the tight and discrete ones of the maximal approach. In some cases, N-cliques can be found that have a property that is probably undesirable for some purposes: it is possible for members of N-cliques to be connected by actors who are not, themselves, members of the clique. The sort of a restriction has the impact of forcing all paths among members of an n-clique to happen by way of different members of the n-clique [3]. This approach is the N-Clan. The k-plex approach would seem to have a considerable amount in common with the n-clique approach, yet k-plex analysis regularly gives a significant distinctive picture of the substructures of a graph.

An anomaly is defined as an unusual action showing an alternate behavior than others exhibit in a similar structure. The term additionally called an outlier, abnormality or special case, has been defined from numerous points of view by various authors. Anomaly detection is additionally viewed as similar to novelty detection in which already unobserved novel patterns in the data are detected [4]. They may at first appear to be same however in novelty detection upon the confirmation of new points they are for the most part incorporated into the model of normal behavior. The presence of anomalies in our data poses numerous problems which should be tackled deliberately. For instance, some kind of malicious users may construct a set of false identities and utilize them to communicate with an extensive random set of innocent users. Henceforth, detection of these anomalous activities in a network is a major concern as their presence may lead to heavy losses [5]. The anomalies are classified into chiefly three categories based upon the nature and scope of anomalies. Point anomalies, additionally known as global anomalies are found if a data object (i.e. a point) demonstrates an alternate behavior than that of the rest of the data. Also known as conditional anomalies, contextual anomalies are available in the data set if the data object deviates essentially with respect to a particular context [6]. Collective anomalies are encountered at whatever point a collection of data objects

all in all depicts an alternate behavior than others, though the individual data objects may not be anomalous. Recently, another sort of anomaly, called horizontal anomaly has evolved in social networks which depict the presence of anomalies based upon the diverse sources of data available [7]. A dynamic anomaly exists with respect to past network behavior in which changes happen in the network with the passage of time. A static anomaly occurs with respect to remainder of the network ignoring the time factor. Labeled anomalies are identified with both structure of the network and the information gathered from vertex or edge attributes. Unlabeled anomalies are connected just to the network structure. No trait of a node or an edge is contemplated.

METHODOLOGY

This work is based on to detect the fake Facebook accounts on the basis of activities of the users. In the existing technique, the formulas are applied on the basis of strength and no of accounts joined. The formulas applied will classify the accounts into fake, moderate and genuine. In this work, the improvement in the existing system is done in which bloom filter is applied which classify the data into fake and genuine accounts. Strength of nodes is calculated with summation of indegree and outdegree of nodes.

Calculation of trust score:

$$\text{Trust Score} = F_a / Z_a$$

Where, F_a is the number of friend request accepted by the nodes in the network which is being sent by the node „a“ and Z_a is the number of friend requests made by the node „a“. The symbol „a“ represent any node in network.

Bloom filter is a type of probabilistic data structure which searches an item which is definitely in collection or not at all in collection [14]. Bloom filter is applied to both trust and strength of nodes in this work. Bloom filter finds a best value from all possible values after execution of no of iterations. The best value selected by bloom filter is used for classification of accounts as genuine or fake accounts class. This way bloom filter combination with trust and strength of nodes provides a more accurate results to classify accounts.

Algorithm 1:

Input: Facebook Account information

Output: Classification of accounts into two classes

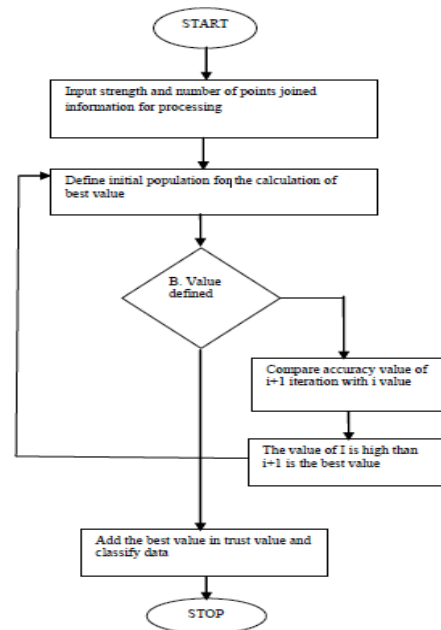
1. A=Strength of account
2. B=Number of points joined
3. Calculate best value ()
4. B=Random point selected from the dataset
5. For (i=0; i=n; i++)
6. Calculate best value $F=F(i)/F''(i)$
7. end
8. If($\text{best}(i) > \text{best}(i+1)$)
9. Best value= $\text{best}(i)$
10. Else
11. Repeat step 5 to 9

12. Calculate accuracy

13. Accuracy=number of accounts classified /Total accounts

14. STOP

This algorithm is developed to classify the malicious accounts from online social networks. The code is implemented in MATLAB to obtain desired results after applying above calculations of trust score and strength. Best value is calculated from both trust score and strength to classify accounts into two classes that is fake and Genuine.



RESULTS AND DISCUSSION

The MATLAB is the tool which is used to perform simulation of proposed and existing models. The technique will be proposed which will be based on the bloom filter technique. In the technique of bloom filter the categorization the users into the two classed means the fake and genuine classes.

Node	Trust Score	Strength	Classification
Node 1	0.25	4	Genuine
Node 2	1	4	Genuine
Node 3	0.5	4	Genuine
Node 4	1	5	Fake
Node 5	1	6	Genuine
Node 6	1	5	Fake
Node 7	1	3	Genuine
Node 8	0.8888	6	Genuine
Node 9	0.5	4	Genuine
Node 10	1	6	Genuine

Fig. 2 Interface of implementation

As shown in the figure 2, the interface is designed which classify the Facebook accounts as fake or genuine.

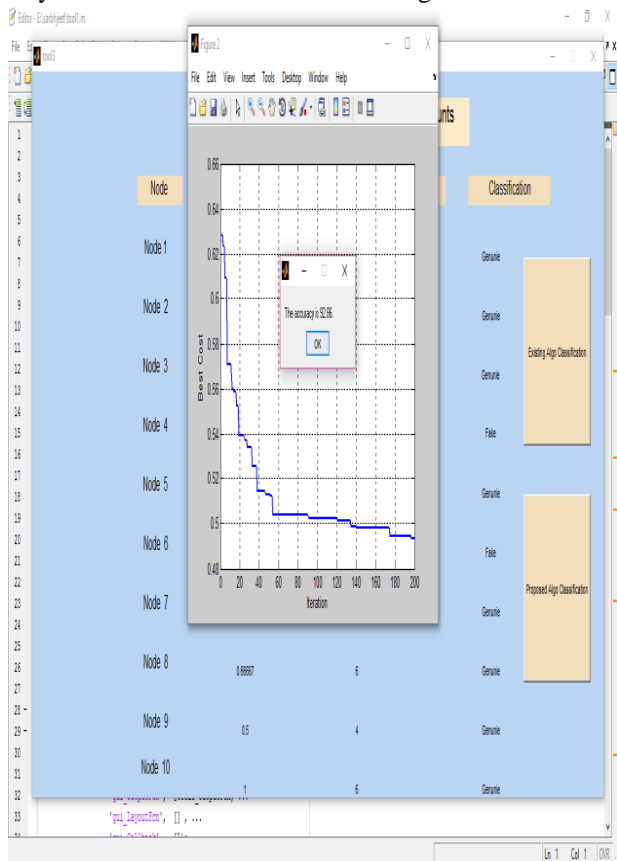


Fig. 3 Final result of proposed work

As shown in figure 3, final result of proposed work is represented by this interface.

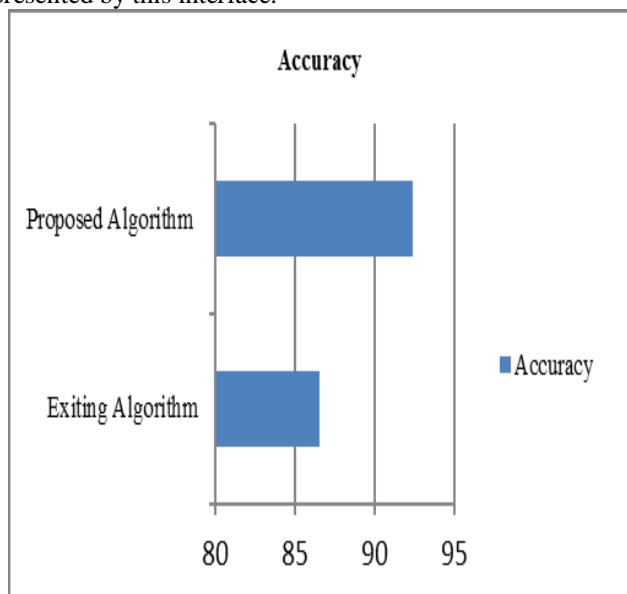


Fig. 4 Accuracy Comparison

As shown in figure 4, the accuracy of the proposed and existing algorithm is being compared and it is being analyzed that accuracy of proposed algorithm is increased to 92 percent from 86 percent.

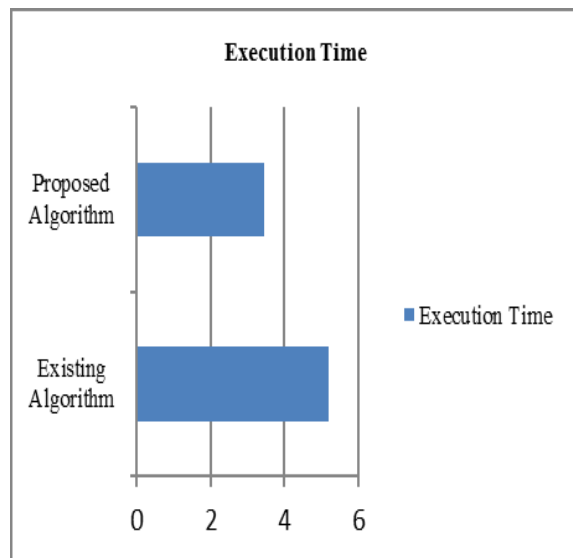


Fig. 5 Execution Time

As shown in figure 5, the execution time of the proposed algorithm is reduced as compared to existing algorithm.

CONCLUSION

In this work, it is being concluded that in the existing algorithm the social network accounts are classified into three classes which are fake, moderate and genuine. In this work, bloom filter is applied which can classify the accounts into two classes which are fake and genuine. The simulation is being performed in MATLAB and it is being analyzed that execution time is reduced and accuracy is increased as a result. In this work we used standard bloom filter. In future other parameters like use of scalable bloom filters can be used for more authentications.

REFERENCES

- [1] Abdolazim Rezaei, Zarinah Mohd Kasirun, Vala Ali Rohani, Touraj Khodadadi, "Anomaly Detection in Online Social Networks Using Structure-Based Technique", The 8th International Conference for Internet Technology and Secured Transactions (ICITST-2013).
- [2] Shota Saito, Ryota Tomioka, Kenji Yamanishi, "Early detection of persistent topics in social networks", 2015, Soc. Netw. Anal. Min, pp. 5:19
- [3] Anita Zakrzewska and David A. Bader, "A Dynamic Algorithm for Local Community Detection in Graphs", 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining.
- [4] Renjun Hu, Charu C. Aggarwal, Shuai Ma, and Jinpeng Huai, "An Embedding Approach to Anomaly Detection", 2016, IEEE
- [5] Ravneet Kaur, Sarbjeet Singh, "Detecting Anomalies in Online Social Networks using Graph Metrics", 2015, IEEE
- [6] P. Kayalvizhi, C. Anoor Selvi, "Detecting Dynamic Topics in Social Network Using Citation based Anomaly Detection", 2015, IEEE Sponsored 9th International Conference on Intelligent Systems and Control (ISCO)

- [7] Evangelos E. Papalexakis, Alex Beutel, Peter Steenkiste, "Network Anomaly Detection using Co-clustering", 2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining
- [8] Weiling Chen, Chai Kiat Yeo, Chiew Tong Lau, Bu Sung Lee, "Behavior Deviation: An Anomaly Detection View of Rumor Preemption", 2016, IEEE
- [9] CuneytGurcanAkcora, Barbara Carminati, Elena Ferrari, Murat Kantarcioglu, "Detecting anomalies in social network data consumption", 2014, Soc. Network Analysis Min
- [10] Hamid Alipour, Youssif B. Al-Nashif, Pratik Satam and Salim Hariri, "Wireless Anomaly Detection based on IEEE 802.11 Behavior Analysis", 2015, IEEE
- [11] Flora Amato, Giovanni Cozzolino, AntoninoMazzeo and Sara Romano, "Detecting anomalies inTwitter stream for public security issues", 2016, IEEE
- [12] William Eberle, Lawrence Holder, "Streaming Data Analytics for Anomalies in Graphs", 2015, IEEE
- [13] David Savage, Xiuzhen Jenny Zhang, Xinghuo Yu, Qingmai Wang, "Anomaly Detection in Online Social Networks", 2014, Social Networks, Volume 39, pp. 62-70, ISSN: 0378-8733
- [14] ShahabeddinGeravand, MahmoodAhmadi "Bloom filter application in network security: A-state-of-the-art-of-survey" 2013 Elsevier journal, computer networks, ScienceDirect.

research focuses in the areas of Network Security, Data Mining, Cloud Computing and Big data analytics.

Authors Profile

HEMASRI ATTURU, received Bachelor of Computer Science degree from Simhapuri University, Nellore in the year of 2013-2016. Pursuing Master of Computer Applications from Sri Venkateswara University, Tirupati in the year of 2016-2019. Research interest in the field of Computer Science in the area of Network Security, Cloud Computing, Networking and Software Engineering.



Dr. Mooramreddy Sreedevi, She is Working as a Senior Assistant Professor in the Dept. of Computer Science, S.V.University, Tirupati since 2007. She obtained her Ph.D. Computer Science from S.V.University, Tirupati. She acted as a Deputy Warden for women for 4 years and also acted as a Lady Representative for 2years in SVU Teachers Association, S.V.University, Tirupati. She Published 40 research papers in UGC reputed journals, Participated in 32 International Conferences and 46 National conferences. She acted as a Resource person for different universities. Her current

