IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727,
PP 01-04
www.iosrjournals.org

A Pragmatic Analysis to Detect Escalating Implicit Behavior in Online Social Websites

Pran Dev[1], Jyoti[2], Dr. Sanjeev Dhawan[3], Dr. Kulvinder Singh[4]
[1, 2, 3, 4](*Department of Computer Science & Engineering, University Institute of Engineering & Technology, Kurukshetra University, Kurukshetra, Haryana, India)*

ABSTRACT : *Now a days, popularity of Internet has been increasing tremendously. Also official work, bill payments, shopping etc. has become online. So, in this era Internet is a need rather than mere interest. And the new generation is using Internet very passionately by means of mobile phones, I-Pads, laptops etc. As a result of it, social interactions among people has been transfiguring online. An increasing number of users making social interactions online give a resilience enclave to social networking sites. With increase in size of social structures, the user with off beam intents also increases. So, to deal with the problem of escalating implicit behavior, a strong analysis of the nodal behavior in online social sites is required. In this paper, an attempt has been made to demonstrate the existing approaches incorporating detection of anomalous users' behavior in online social networks like Facebook, Twitter, Myspace, LiveJournal etc. Digging into a comparative study of existing methods, efforts are made to discuss flaws and pros in these approaches, which can help in the exposure of anomalous users' behavior in online social networks.*

Keywords - *Content Based Analysis, Implicit Behavior, Intent Recognition, Link Based Analysis, Social Network Analysis*

I. INTRODUCTION

Online social networks like Facebook, Myspace allow users to share pictures, videos, texts etc. and users can share their moods (a recent application of Facebook) and they also let users establish the social relationships which are vanishing due to less social interactions. So, there are many reasons for the enormous expansion of online social websites and the exponential growth in number of users provokes researches to study the structural properties of online social networks. Online social networks are widely used for information propagation, this information propagation must be controlled as there are both correct information and rumors. Other problems that occur in online social networks include unwanted communication like spams, unwanted friends, unsolicited messages etc. Also there is a problem of finding legitimate and malicious users. In online social networks, there exists communities; a user can be member of more than one community. So, the problem of finding overlapping communities should also be taken into consideration for the betterment of OSNs. Online social network analysis is done for the evaluation and improvement of the properties of OSNs. Social network analysis is basically of two types:

1 Content Based
2 Link Based

In content based social network analysis, the analysis is done on the basis of contents available in social networks like messages, wall posts, tweets etc. And in Link based analysis, the structure of social networks is analyzed in form of online social networks graphs, random link analysis, positive-negative link prediction etc.

Considering all these problems in online social networks, in this paper, a study of existing approaches regarding above mentioned flaws is carried out which can help in improving the quality, trust and security in OSNs. Previously, a lot of work is done on both content based and link based approaches of SNA, here the study of both these methodologies is carried out.

In this paper, Section I presents brief introduction about the online social network and analysis of OSNs. Section II describes the previous related work carried out in the field of OSNs, link based as well as content based and discusses the existing methodologies to detect implicit behavior, to find spams and other approaches to deal with several online social networks' problems and also discusses the pros and cons of the existing techniques in brief. And Finally Section III concludes the paper and provides future scope that can be useful for the betterment of online social networks by improving the existing approaches.

II. RELATED WORK

The Anomalous users' behavior can be detected in online social networks by analyzing the flow of content i.e. information in network structure. Also by studying both the usual and unusual activities, particular properties of the users can be identified on the basis of which anomalous users can be detected. The size of social networks is so large and analyzing such large datasets is quite a hard work to carry out, which is attracting researchers. In this paper, work carried out on the problems like anomaly detection, finding hidden relationships, malicious users' behavior and detection of spam nodes is analyzed in detail. To detect regular behavior, Lahiri and Tanya [1] projected a new mining problem of finding sub-graphs in dynamic social networks on the basis of periodicity. They explored the computational paradigms of the problem. They anticipated a real-world, effectual and mountable algorithm to find sub-graphs that precedes deficient periodicity into explanation. Then Clifford Weinstein *et al.* [2] labelled a tactic for "modeling, detection, and tracking of terrorist groups and their objectives based on multimedia data". They concentrated on demonstrating and recreation of terrorist attacks using the data available of the previous attacks. Next, a two-phase strategy was assumed by Lei Tang *et al.* [3], to categorize the hidden relationships pooled across scopes in multi-dimensional networks. They mined operational structures from multiple dimensions of the network and then united them all to discover the adhesive community architecture from users. A "context-aware" content examination agenda was proposed by Cheong and Lee [4] to mine dormant properties from tweets in Twitter. In accumulation, they assimilated an unsubstantiated "Self-organizing Feature Map" [4] as a machine learning approach. They delivered an approach to determine users' customs and feelings when subsidizing to prevalent subjects of conversation on Twitter. Then, Yassine and Hajj [5], proposed an innovative structure for portraying emotive relations in social networks, and then exhausting these arrivals to discriminate friends from acquaintances. The objective is to excerpt the emotive content of texts in online social networks. For this tenacity, text mining practices are achieved on remarks regained from a social network. They vacant a fresh standpoint for revising friendship associations and emotions' expression in online social networks; it chiefly uses the k-means clustering algorithm. This work was sustained by Gianluca Stringhini *et al.* [6], they investigated the degree to which spam has increased in heterogeneous online social networks. Further accurately, they studied the functioning of spammers in online social sites. To accumulate the information about spammers' doings, they formed a big and varied class of "honey-profiles" [6]. Built on the exploration of this behavior, methods are settled to perceive spammers in social networks, and their messages are collected in large spam drives. They also disclosed how this procedure aid to identify spam profiles even when they do not interact a honey-profile. Moreover, Hongyu Gao *et al*. [7] presented a preliminary learning to calculate and portray spam activities propelled using interpretations on online social websites. They studied a outsized mysterious dataset of not synchronous "wall" messages communicated among users in Facebook. Their interpretations revealed that when accustomed to the resident time of the source, spamming governs authentic wall post motion in the initial morning time, when regular nodes are inactive or sleeping. The pragmatic assessment was then done by Yan and Wang [8], they emphasized straight on the vertex and link characteristics in an elongated time stamp. The link formation and omission method was probed to build sprouting behavior patterns in online social websites. Primarily, the regular varying behaviors idea was offered, and the properties of normal behaving nodes were explored by a procedure based on "Stochastic process and Markov model" [8]. Then, a spectrum based detection framework was given by Xiaowei Ying *et al.* [9], to notice the committers of attacks. They concentrated on "Random Link Attacks" [9] in which the mischievous user makes numerous deceitful identities and relationships, later by using these fake relationships the spam attack is expended in larger region. They specified how to categorize attackers by detecting their disseminations in the spectral space. They offered a fresh agenda that adventures the spectral space of fundamental network structure to recognize scams or spams. Another security mechanism was given to identify spam activities in twitter messages by Kristofer Beck [10], he looked at discovering spammer behavior of the Twitter social network. He specified that malicious nodes use definite keywords to lure a twitter user to click on any unwanted link. This link could clue them to a mischievous web page. This twitter spam detection was further studied and experimented by Charles Perez *et al.* [11], by making use of SPOT. It goals to suggest a structure to evaluate disbelieving behavior on Twitter. They offered automation established for finding doubtful users' profiles on Twitter network. A naïve method was then given by Zahid Halim *et al.* [12] by means of space based and time based features. To detect the unusual behaving users they used the spatio-temporal characteristics i.e. nodes which communicate regularly and are involved in illegal activities used semantic analysis, as a basis of deducing draws between users. Then a step toward the new experiment of inspecting coupled behavior was carried out by Longbing Cao [13], he deliberated the problem of Coupled Behavior Analysis. "A Coupled Hidden Markov Model" [13], based on this, a method is demonstrated to archetype and

discovers irregular community-based exchange behaviors. A new strategy was formulated by Dora Erdos *et al.* [14], they established a common procedure to treat with the problem of restructuring graph from locality data. They highlighted on rebuilding the concealed binary matrix that specifies the manifestation or nonappearance of associations amid different users. Then investigation on content produced by various users on online social sites was done for identifying spam behavior by Enhua Tan *et al.* [15], they claimed that spammers often enclose famous keywords or merely duplicates current objects from the internet, trying to restrict information exposure. Their examination displays that the spammers reveal exclusive non-textual patterns. Based on these non-textual properties, they presented that by means of numerous classification approaches that a great recognition rate could be attained offline. After that there was another statement by Dae-Ha Park *et al.* [16] about detecting spams in online sites using feature extraction and Bayesian classifier. They improved the present renowned classification procedures such as Bayesian network classifiers (BNCs) to adapt for SNS features. Further, an in-depth investigation was carried out by Pasquale De Meo *et al.* [17], they deliberated three prevalent platforms, Flickr, Delicious and StumbleUpon, and, by merging practices from social network examination with performances from semantic study, they categorized the tagging behavior as well as the inclination to generate friendship relationships of the nodes of these platforms. Then, a new idea was brought in light by Manuel Egele et *al.* [18] for anomaly detection. They offered a fresh tactic to perceive compromised user accounts in social networks, their method makes use of an arrangement of statistical modeling and abnormality uncovering to recognize records that practice an abrupt alteration in behavior. They settled a tool, called COMPA, COMPA was capable of recognizing cooperated accounts on both social networks with great exactness. David Mandell Freeman [19] termed a class of structures by making use of a Naive Bayes classification algorithm to discover accounts which are fake users. Zejia Chen *et al.* [20] investigated the problem of spammer detection from the perspective of users' behavior. They proposed a cascading framework for detection of spammers in online social networks called CWB-SPAM [20]. Furthermore, Hongzhi Yin *et al.* [21] Focused on analyzing users' behavior in online social networks and designed a "latent class statistical model" [21], to find the relationships between different users and to analyses their behavior. They took two factors in consideration for their purposed work, one is user deep-down interest as an implicit feature and other is time based context as explicit feature. And then, M. Sahlabadi [22] proposed a technique to identify the anomalous user behavior, the first step is finding characteristics of normal user behavioral pattern and the second step is then detecting abnormal behavior by measuring the deviation from the characteristics of normal behavior user.

III. CONCLUSION AND FUTURE SCOPE

With advent in interactions between the entities in the network, the behavior of users in online social network has become an enigma. Discovering countless communities in the network is of vital importance. There is a dire need of an affirmative study to have a deep understanding of social networks and detecting hidden patterns. Despite of aiming only on global structure of user interactions, an in-depth scrutiny is paid in this paper. Behavior analysis plays an imperative role in perceiving hidden relationships, node associations, and exposure of spam activities in online social networks, which, in turn, aids in exterminating the security issues, making social networking more contented. In this paper, an attempt has been made to study the existing patterns of interactions, to analyze the shortcomings and pros of the already existing techniques for detection of hidden relationships, and to give an epitome of the related work done in this area. Techniques discussed in this paper can be enhanced further by using genetic algorithm, fuzzy systems etc., which in turn, can help in making online social networks more prominent and secure for the users to share their personal data with their friends or circles.

REFERENCES

[1] M. Lahiri and T. Y. Berger-Wolf, Mining Periodic Behavior in Dynamic Social Networks, In *Eighth IEEE International Conference on Data Mining, IEEE*, 2008, 373-382.
[2] C. Weinstein, W. Campbell, B. Delaney and G. O'Leary, Modeling and Detection Techniques for Counter-Terror Social Network Analysis and Intent Recognition, In *Aerospace conference, IEEE*, 2009, 1-16.
[3] L. Tang, X. Wang and H. Liu, Uncovering Groups Via Heterogeneous Interaction Analysis, In *Ninth International Conference on Data Mining, IEEE*, 2009, 503-512.
[4] M. Cheong and V. Lee, A Study on Detecting Patterns in Twitter Intra-Topic User and Message Clustering, In *20th International Conference on Pattern Recognition (ICPR), IEEE*, 2010, 3125-3128.
[5] M. Yassine and H. Hajj, A Framework for Emotion Mining from Text in Online Social Networks, In *International Conference on Data Mining Workshops (ICDMW), IEEE*, 2010, 1136-1142.

[6] G. Stringhini, C. Kruegel and G. Vigna, Detecting Spammers on Social Networks, In *Proceedings of the 26th Annual Computer Security Applications Conference*, ACM, 2010, 1-9.
[7] H. Gao, J. Hu, C. Wilson, Z. Li, Y. Chen and B. Y. Zhao, Detecting and Characterizing Social Spam Campaigns, In *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*, ACM, 2010, 35-47.
[8] L. Yan and J. Wang, Extracting Regular Behaviors from Social Media Networks, In *Third International Conference on Multimedia Information Networking and Security (MINES), IEEE*, 2011, 613-617.
[9] X. Ying, X. Wu and D. Barbará, Spectrum Based Fraud Detection in Social Networks, In *27th International Conference on Data Engineering (ICDE), IEEE*, 2011, 912-923.
[10] K. Beck, Analyzing Tweets to Identify Malicious Messages, In *International Conference on Electro/Information Technology (EIT), IEEE*, 2011, 1-5.
[11] C. Perez, M. Lemercier, B. Birregah and A. Corpel, SPOT 1.0: Scoring Suspicious Profiles on Twitter, In *International Conference on Advances in Social Networks Analysis and Mining (ASONAM), IEEE*, 2011, 377-381.
[12] Z. Halim, M. M. Gul, R. Baig, S. U. Rehman and F. Naz, Malicious Users' Circle Detection in Social Network Based on Spatio-Temporal Co-Occurrence, In *International Conference on Computer Networks and Information Technology (ICCNIT), IEEE*, 2011, 35-39.
[13] L. Cao, Y. Ou and P. S. Yu, Coupled Behavior Analysis with Applications, *IEEE Transactions on Knowledge and Data Engineering*, Vol. 24(8), 2012, 1378-1392.
[14] D. Erdös, R. Gemulla and E. Terzi, Reconstructing Graphs from Neighborhood Data, In *ICDM*, 2012, 231-240.
[15] E. Tan, L. Guo, S. Chen, X. Zhang and Y. Zhao, Spammer Behavior Analysis and Detection in User Generated Content on Social Networks, In *32nd International Conference on Distributed Computing Systems (ICDCS), IEEE*, 2012, 305-314.
[16] D-H. Park, E. Cho and B-W. On, Social Spam Discovery using Bayesian Network Classifiers based on Feature Extractions, In *12th IEEE International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom), IEEE*, 2013, 1808-1811.
[17] P. Meo, E. Ferrara, F. Abel, L. Aroyo and G. Houben, Analyzing User Behavior across Social Sharing Environments, In *Transactions on Intelligent Systems and Technology (TIST)*, ACM, 2013.
[18] M. Egele, G. Stringhini, C. Kruegel and G. Vigna, COMPA: Detecting Compromised Accounts on Social Networks, In *NDSS*, 2013.
[19] D. M. Freeman, Using Naive Bayes to Detect Spammy Names in Social Networks, In *Proceedings of the 2013 ACM workshop on Artificial intelligence and security*, ACM, 2013, 3-12.
[20] Z. Chen, J. Yang and J. H. Wang, A Cascading Framework for Uncovering Spammers in Social Networks, In *Networking Conference, IFIP*, IEEE, 2014, 1-9.
[21] H. Yin, B. Cui, L. Chen, Z. Hu and Z. Huang, A Temporal Context-Aware Model for User Behavior Modeling in Social Media Systems, In *Proceedings of the 2014 ACM SIGMOD international conference on Management of data*, ACM, 2014, 1543-1554.
[22] M. Sahlabadi, R. C. Muniyandi and Z. Shukur, Detecting Abnormal Behavior in Social Network Websites by Using a Process Mining Technique, In *Journal of Computer Science*, 10(3), 2014, 393-402.