International Journal of Computer Science and Engineering Communications,
Volume.5, Issue.3 (2017): Page.1600-1605
www.ijcsec.com

Graph Based Real Sybil Detection in OSNs

Shansa P[1] and Vigilkumar V.V[2]
[1]PG Scholar, [2]Assistant Professor
[1,2]Department of Information Technology
[1,2]Govt.Engineering College Idukki, India
shansashanmughan4@gmail.com[1],vigil.lit@gmail.com[2]

Abstract: *In current generation, the rapid growth of Online Social Networks (OSNs) have made far reaching effect on everyone's social life. This increase in popularity, usage and anonymous nature of OSNs exposed the possibility of being attacked. In Sybil attack, a fake user can create massive amount of fake identities towards the target OSNs for unfairly increasing their influence. These Sybils performs the distribution of malwares, spams, bad products reviews and private data collection. Recently, there exist different schemes for detect and prevent the challenging Sybil attacks. This paper review some of those works that leverages the social graph structure and make a comparative study to identify their relevance in detection of Sybil identity in OSNs.*

Keywords *- Online Social Network, Sybil attack, Sybil detection*

I. Introduction

The social networking concepts were first investigated during the year 1960s, which become an indispensable part of today's human lives. Among the social networking sites, the largest OSNs such as Facebook and Twitter has 1.3 billion registered users. This OSNs allow its users to create profiles pages containing personal details, making, connecting and keeping in contact with friends and often connect with a lot of strangers. Due to this popularity and the open nature OSNs are more vulnerable or susceptible to different type of attacks. One of this attack is Sybil attack, where malicious user create multiple accounts to gain knowledge about a lot of sensitive personal information from user. Existence of Sybil attack in OSNs not only effect it's users but also create negative impact in OSN's advertising and marketing fields.

To prevent the Sybil attack, or to limit the impact of the Sybil attack in OSN, some of the earlier methods were IP Address Tracking, Content Analysis, Account Activity Statistics and User Complaints. Advanced techniques are required for prune the Sybils presented in network. In order to isolate fakes in OSN, some works leverages the social graph structure. The graph structure based schemes that use social network information are more successful comparing to the previously mentioned methods to defend against Sybil attacks. Most of these social graph structure based detection techniques [1], [2], [5], [7] are relay on the key assumption that, Sybil accounts have the difficulty to befriending many real users. They form limited attack edges. There exist loose connection between the real region and Sybil region. Fig. 1 shows the representation of different nodes and attack edges between them in SybilInfer [5].

Some detection techniques focus on community structure of Sybil nodes. Sybil identifications were done by using different community detection methods. Sybils can escape from this community techniques by sending large number of friend request towards the legitimate region. Studies shows that real users are careless about friend request acceptance, not all Sybils try to form community structure and there exist advance technology for create more realistic profile. These are some of the reasons for decreasing the relevance of community detection schemes. Some graph based works focus on the positive trust relationship and some on distrust relationship.

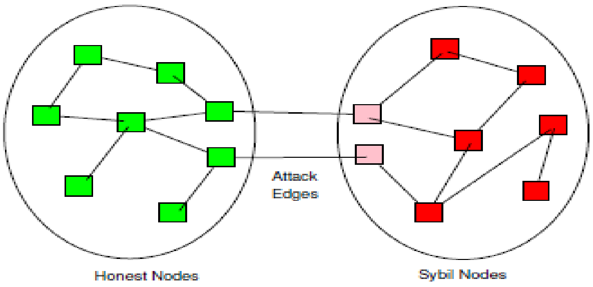

Fig. 1. Attack edges between Honest nodes and Sybil nodes [5]

The rest of this paper is organized as follows. Section II covers the literature survey of some social graph structure leveraging detection schemes and comparative study of these mechanism is carried out in Section III for identifying their relevance in Online Social Network for real Sybil detection. Finally, conclusion about the survey is presented in Section IV.

II. Literature Survey

H. Yu *et al*. propose SybilGuard [1], protocol for detecting Sybil attacks in social networks. This popular decentralized protocol reduce the influences of Sybil nodes presented in network. In SybilGuard, the entire functionalities are performed with respect to a node and is completely decentralized. SybilGuard is actually based on the social network among the users and the links between two user identities indicates a trust relationship which is human-established. Sybil guard leverages special kind of random walks. Each node in network perform random routes. The verifier's random route accepts the suspect when the random routes intersect.

Sybil Guard relies on the social network properties:

(i) The real region of the social network is fast mixing, and

(ii) Malicious user identities may create many nodes but can for relatively less number of attack edges.

Sybil guard protocol can detect only one Sybil node at a time. It suffers from higher false negatives. In order to find out the Sybil region, entire nodes presented in the network need to be properly examine. System fails to correctly distinguish honest and Sybil communities.

H. Yu *et al.* then propose another protocol called SybilLimit [2], an advanced version of SybilGuard. This protocol is also decentralized. Sybil Limit accepts only O(log n) Sybils per attack edge from sybil region to real and are on the same assumptions that are considered in SybilGuard. SybilLimit exploit the fast mixing property of social network. Experiments for a million number node system shows that, the amount of Sybil nodes accepted by Sybillimit is reduce to 200 times comparing to SybilGuard. This reduction is because of a short random walk, which is likely to stay within the set of honest node. In Sybil Limit, all the suspect nodes in the social graph are tested for identifying the Sybil nodes. The protocol is based on unrealistic assumption regarding the number of honest nodes in networks.

Another approach, SybilInfer [5] a centralized Sybil defense algorithm, leverages the Bayesian inference approach. It label each nodes in network as real or Sybil by assigning a Sybil probability. That probability represents the degree of certainty of each node in the network to be a Sybil. Based on this degree of Sybil certainty, the ranking will be provided. The algorithm takes the network graph and a known single honest node as input. SybilInfer obtain lower false negatives, but at a higher cost of computation overhead. But Sybil Infer can handled only a network that containing 30K nodes. When comparing the network size with regular online social network, a network with 30k nodes is really smaller .Sybil Infer can handle only smaller networks.

SybilDefender [3], another detection scheme for Sybil detection by W. Wei *et al*. This system leverages the network topologies for defending Sybil attacks in the case of large social networks. It is more efficient and also scalable in large social networks. SybilDefender have two algorithms. One for identifying the Sybil nodes and another for detecting Sybil community in network. At a time this system detects a single group of Sybils. Sybil identification algorithm identifies the individual Sybil nodes. Then those nodes are used to identify the surrounding community of Sybil nodes. Within one run of the algorithm the community of Sybils around an individual Sybil node can be detected. SybilDefender is capable of detecting communities in unsigned graphs.

Q. Cao *et al*. present SybilRank [6], an effective and efficient fake account inference scheme, which allows OSNs to perform ranking among user accounts according to their perceived likelihood of being fake. The assumption considered in this work are social graph is undirected, non-sybil region is well connected and limited number of attack edges are possible. The computational cost of SybilRank does not increase with the number of selected seeds. It attain high accuracy at low computational cost. It discover all the Sybils from an already known seeds. System uses multiple seed selection in order to avoid seed target attack. The multiple seed selection make hard for the Sybil to attack the target seeds. So seed selection errors are limited in this scheme.

N. Tran *et al.* propose SumUp [7], a Sybilresilient vote aggregation system. This system leverages the network of trust among users for defending against Sybil attacks in social network. SumUp uses the adaptive vote flow aggregation technique. This adaptive vote flow aggregation helps to reduce the amount of fake votes casted by adversaries. It limit the fake vote to not more than the count of attack edges as voting feedback. SumUp also limit the power of adversaries voting through user feedback on votes. The assumption considered by the system is Sybils can make few real users as friends.

Integro [8], a scalable hybrid defense system by Y. Boshmaf *et. al*. As the Sybils spreads spams and malwares, both the OSN operator and its users require a fake detection system. Integro incorporate feature extraction with graph structure for Sybil identification. This system predict the victim accounts using a classifier and perform meaningful user ranking in graph structure. Integro integrate the user level activities into social graph structure for pruning Sybil nodes presented in network. The system provides higher ranking for real accounts comparing to fake accounts. OSN operators can take proper actions against those Sybil accounts having low-ranking in the ranking scheme.

Dieudonne Mulamba *et.al* propose SybilRadar [9] which is a powerful Sybil detection architecture based on structural properties of an OSN .It does not depend on the conventional non-realistic assumptions that similar structure based frameworks make. It uses structural properties of the OSN graph using similarity metrics. Without using content- based technique SybilRadar produce much accuracy. Performance of SybilRadar is better even in the OSNs that have tremendously large number of attack edges. The high computational overhead is there in the SybilRadar.

SybilFence [10] is based on the examination that even well-maintained fake user accounts definitely receive a remarkable number of negative feedbacks from user. They consider the rejections to the friend requests. Their key idea is to discount those edges across users which receives feedbacks as negative. Only the negative feedback from users are used for detect the Sybils in OSNs, thereby reducing the influence of Sybils social edges. System directly apply the negative distrust relationship, which can be easily manipulate by the Sybil nodes using another Sybils who are granted to accept those requests.

VoteTrust [4] presented by Z. Yang *et al*. models the interactions of friend invitation among users as a signed directed graph. Friend invitation graph with an edge that directed from the sender node to the receiver node and a value indicates whether a friend request is accepted or not. This work shows that Sybils can befriend huge number of real users by sending a large amount friend requests, invalidating the unrealistic assumption behind social-graph-based detection. Actual difficulty of Sybils is to require real users to befriend them first or to accept them with a high probability. It uses two main mechanisms to identify Sybils over the-graph: a voting-based real individual Sybil detection and a Sybil community identification. Voting-based detection is used to discover the presence of individual Sybils using the trust relationship. Sybil community detection find other colluding Sybils around the identified Sybils. Votetrust focus on both acceptance and rejection behavior.

III. Comparative study

Based on the literature survey, a comparative study on different graph based Sybil detections in OSNs is done and are shown in Table.1. Detailed classification of each research works including their main advantages and disadvantages are included and besides emphasizing on defining the merits and demerits, the main techniques are pointed out.

Table.1 Comparative Study

SYSTEM	METHOD	ADVANTAGES	DISADVANTAGES
SybilGuard [1]	Decentralized protocol based on the social network	Leverages the social graph structure. Reduces Sybil attack impact in social networks.	Cannot detect more than one Sybil at a time. Detect the region of Sybils, all nodes need to be examined. Approach fails to distinguish between honest and Sybil communities. Large number of sybil nodes are allowed to be accepted.
SybilLimit [2]	Decentralized approach using topological feature.	Leverages the social graph structure. Limits the number of sybil nodes.	Cannot detect more than one Sybil at a time. Detect the region of Sybils, all nodes need to be examined. Unrealistic assumptions about the honest nodes number in the network.
SybilInfer [5]	Topological feature using Bayesian inference approach	Single known good node. Sybil probability assigned to nodes using Bayesian Inference. Low false negatives. Algorithm output by a degree of certainty.	High computation cost. Handled networks only with up to 30K nodes.
SybilDefender [3]	Community detection by leverages the network topologies	Detect sybil community. Minimum number of random walks. Efficient and scalable to large social networks.	Fake accounts establish sparse connectivity among themselves. Cannot detect communities in signed graphs.

SumUp [7]	Adaptive vote flow technique.	Leverage positive trust relationship. Limit bogus votes cast by adversaries to no more than the number of attack edges.	Assumption that Sybils can befriend only few reals. High computational requirements.
SybilRank [6]	Social Graph properties to rank users	Leverage the positive trust relationship. Leverages its efficient support for multiple trust seeds to reduce the false positives. Low false positive ranked list.	Social graph is undirected. Ineffective when fakes infiltrate the OSN by befriending a large number of real users.
Integro [8]	Integrates user activities into graph structures.	Most real accounts rank higher than fakes. Victim accounts knowledge help to handle sybil even if they befriending a large number of real users.	OSN users frequently leave their profiles incomplete. Use misleading information purposefully which are often incomplete, inaccurate or raise privacy concerns.
SybilRadar [9]	Structural properties of the OSN graph using similarity metrics	Produces similar detection accuracy without using any content- based techniques. Does not rely on the traditional nonrealistic assumptions.	Computational overhead.
SybilFence [10]	Social-graph based Sybil detection using rejection	Discount the social edges on users that have received negative feedback.	Negative distrust relationships among users is directly applied.
VoteTrust [4]	Friend Invitation Graph	Only focuses on the friend invitation behavior. Detects Sybils that get more rejections than acceptances from real users, irrespective of the number of victims.	Cannot handle miscreants selling legitimate accounts.

IV. Conclusion

In recent years, Sybil attack presented in online social networks become more crucial and very powerful. Among different detection techniques, the schemes that are concentrated on the social network structure for defending Sybil attack draws more and more attention. Some of those techniques are discussed in this survey. In this paper, we studied and analyze the different way to identify Sybil accounts in online social networks using graph structure. Most of the detection schemes that analyze the graph structure of OSNs are based on unrealistic assumptions. VoteTrust system only consider the actual Sybil difficulty and focus on both acceptance and rejection behavior using friend invitation graph in Sybil detection. But the VoteTrust cannot handle attackers who buy friends from miscreants.

REFERENCES

[1] H. Yu, M. Kaminsky, P. B. Gibbons, and A. Flaxman , Sybilguard: Defending against Sybil attacks via social networks, *IEEE/ACM Transactions On Networking*, Vol. 16, No. 3, June 2008.

[2] H. Yu, P. B. Gibbons, M. Kaminsky, and F. Xiao, Sybillimit: A near-optimal social network defense against sybil attacks. *IEEE/ACM Transactions On Netw*orking, June 2010.

[3] W. Wei, F. Xu, C. C. Tan, and Q. Li, Sybildefender: Defend against sybil attacks in large social networks, IEEE Transactions On Parallel and Distributed Systems, 2013.

[4] J. Xue, Z. Yang, X. Yang, X. Wang, L. Chen, and Y. Dai,Votetrust: Leveraging friend invitation graph to defend against social network sybils, *IEEE Transaction*, June 2016.

[5] G. Danezis and P. Mittal, SybilInfer: Detecting sybil nodes using social networks,*in Proc. Netw. Distrib. Syst. Security*, 2009.

[6] Q. Cao, M. Sirivianos, X. Yang, and T. Pregueiro, Aiding the detection of fake accounts in large scale social online services, *in Proc. 9th USENIX Conf. Netw. Syst. Des. Implementation*, p. 15, 2012

[7] N. Tran, B. Min, J. Li, and L. Subramanian, Sybil resilient online content voting, *in Proc. of NSDI*, 2009

[8]Y. Boshmaf, D. Logothetisy, G. Siganosz, J. L. Jorge Lerax, M. Ripeanu, and K. Beznosov, Integro: Leveraging victim prediction for robust fake account detection in osns, *in Proc. of NDSS,* 2015.

[9] Dieudonne Mulamba, Indrajit Ray, and Indrakshi Ray, SybilRadar: A Graph-Structure Based Framework for Sybil Detection in Online Social Networks,IFIP Advances in Information and Communication Technology, vol 471. Springer,2016.

[10]Qiang Cao Xiaowei Yang,SybilFence: Improving Social-Graph-Based Sybil Defenses with User Negative Feedback, WOSN 2012 on March 14.