International Journal of Applied Engineering Research
ISSN 0973-4562 Volume 10, Number 9 (2015) pp. 24693-24700
© Research India Publications
http://www.ripublication.com

Sybil Defense Mechanisms: A Survey

Akash Suresh[1], Varun Babu[1], Jithin J[1], Jeshwanth Manne[1], Abirami K[1]

*[1] Department of Computer Science, Amrita School of Engineering,
Amrita Vishwa Vidyapeetham University, Ettimadai, Coimbatore-641112, India.*

ABSTRACT

Sybils are fake identities in any network. Sybils cannot be matched to a physical identity. There are several mechanisms to detect Sybil Attacks in networks. In this paper we are discussing Sybil Defense Mechanisms in Social Networks. We have included the working principles of existing Decentralized Sybil Defenses and the assumptions under which these mechanisms operate. We have also gathered information supporting the fact that the assumptions about Social Networks are flawed. Certain drawbacks of each mechanism have also been looked into.

Keywords. Sybils, Attack Edges, Honest Nodes, Benign Nodes, Minimum Cut.

I. INTRODUCTION

A. Sybil Attack

The phrase Sybil Attack originates from the book ‘Sybil’, a case study of a woman diagnosed with several personalities. Forged identities in a network are referred to as Sybils. These manipulate various network parameters. Examples would include the following, few malicious nodes rigging online voting systems, certain systems in a network obtaining access to an unfair amount of shared resources, increasing popularity of certain products on e-commerce [3], [4]. Sybils are also common in anonymous communication and sensor networks. There are various protocols to thwart Sybil Attacks. In this paper, we are discussing existing Sybil Defense Mechanisms via Social Networks.

B. Sybil Attack In Social Networks

Sybil attacks are prominent in Online Social Networks (OSN) such as Facebook and Twitter. In such attacks, an attacker creates multiple identities and manipulates honest nodes into thinking that they too are of authentic origin. Sybils in Facebook work by sending out friend requests in large numbers to benign nodes. It can also be the other way round, i.e., Honest nodes sending requests to Sybil nodes because such nodes appear to be genuine. However, in Twitter the scenario is different. An attacker does not need to accept a friend request to gain access to a profile. Anyone's profile can be accessed by clicking the follow button on Twitter. For example, a fake profile can be created by an attacker to attract benign nodes in the network. Once the attacker makes acquaintance with honest nodes, the node can take control of the benign profiles [5].

II. SYBIL DEFENSES

A. Centralized Or Decentralized Approach

Using centralized systems, where a central authority authorizes the nodes can be effective in controlling and limiting the access of Sybils, but as Social Networks are growing exponentially, monitoring it through a central authority is clearly inefficient. Also, if the central authority is attacked the entire network is affected. There is a single point of failure. This is why decentralized mechanisms are required for the identification of Sybils.

B. Decentralized Mechanisms

There are broadly two categories of decentralized Sybil Defenses-Sybil Detection and Sybil Tolerance. The former method deals with detection of Sybil nodes followed by its removal. The latter category involves reducing impact of Sybils. Sybil Detection methods include-SybilGuard, SybilLimit, SybilRank, SybilInfer and GateKeeper. Sybil Tolerance methods are Ostra and SumUp. All the methods mentioned above are decentralized protocols. These methods provide good experimental results [1], [2].

C. Assumptions Made About Sybils

All the defense mechanisms named above make use of the Social graph structure to identify malicious nodes. The structure of the graph is influenced by the assumptions that are made about Sybil nodes, Honest nodes and the relationship between them. Assumptions made are the following, **i)** it is difficult to establish attack edges, **ii)** existence of two separate clusters of nodes, one containing only Sybils and the other containing only benign nodes connected by a few attack edges. Figure 1(a) shows the structure of one such Social graph.

With such assumptions in place, it is certain that the Social graph will have a minimum cut as there will be only a small number of attack edges connecting the Sybil community to the Honest community [9]. It is this property of the graph that is utilized by the above graph based Sybil Defense mechanisms. Social graphs are fast mixing. If there exists a minimum cut between two nodes the mixing would be slower [2]. A graph that follows this assumption is shown in Figure 1. (a) Figure 1(b) shows a Social graph that resembles a real-world Social Network.

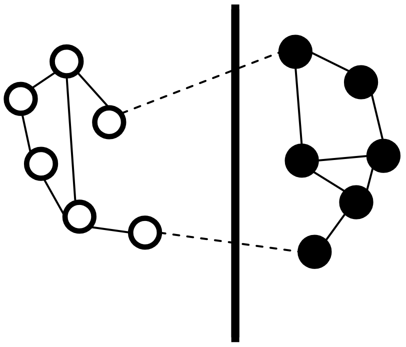

Figure 1. (a) Shows a model that satisfies all the assumptions (discussed in above) made about Social Networks.

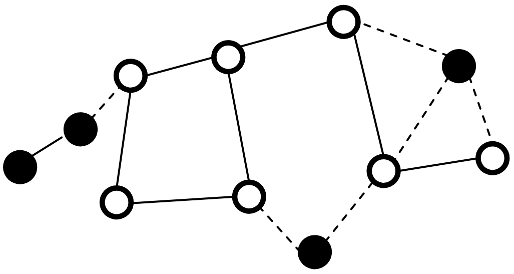

Figure 1. (b) Shows a realistic model proposed [1]

○ Honest Node .... Attack Edge

● Sybil Node

III. NEED FOR AN EFFICIENT SYSTEM Referring to the models based on which the current Sybil

Defenses work as classical and a new model proposed as modern, various drawbacks of the existing methodologies are pointed out. In the real world the Social Networks are a lot different compared to the classical model. In actual Social Networks, **i)** Sybils require only a few edges to disguise itself as a benign node, **ii)** Sybils have links to larger number of benign nodes rather than other Sybils. As a result, distinguishable clusters of Sybil nodes and Honest nodes are not formed [1], [6].

Both these facts contradict the assumptions made while modeling Social Networks. Statistics of the behavior nodes that affect the structure of the graph further confirm this contradiction. The structure of the Social graph depends on **i)** how easily Honest nodes trust Sybils [6], **ii)** 11 Million attack edges for 65,000 Sybils [6], **iii)** 170 attack edges per Sybils [6], **iv)** 50% of the benign nodes trust Sybils [7], **v)** Benign nodes also send request to Sybils [5], [8].

This is why the existing Sybil identification mechanisms are not completely efficient. Also, based on studies from an existing Social Network in China, the ineffectiveness of the classical model is exposed by a group of researchers. Here, the primary observation was that more than seventy percent of the Sybil nodes in their study did not have Social links to other Sybil nodes. Rather, Sybils used snowball sampling techniques to identify users with lots of followers and friends. Many users accept requests from unknown nodes. This way Sybils can disguise as a benign node. For this reason it is an almost impossible feat to identify Sybils with complete accuracy.

These studies also emphasize the need for a more realistic model compared to the classical model as proposed in [1].

On analyzing a sample of Sybils and Non-Sybils (previously verified sample), the group came up with few characteristics that can be used to differentiate between the two. Classification characteristics from their study include,

i) invitation frequency,
ii) fraction of incoming requests accepted,
iii) fraction of outgoing requests accepted,
iv) clustering coefficient (extent to which the nodes are interconnected) [6].

Studies prove invitation frequency and fraction of incoming requests accepted for Sybils are considerably high. On the other hand fraction of outgoing requests accepted and clustering coefficients are low.

IV. SYBIL DETECTION APPROACHES

Next, we are discussing the different Sybil detection approaches and corresponding drawbacks.

A. SybilGuard

Working.

It enables any Benign node (a verifier) to decide whether or not to accept other nodes (suspects). Figure 3. shows a visualization of how a suspect tries establish connection with a verifier. The SybilGuard protocol is based on the intersection of random walks. Depending upon the degree of the verifier and suspect, random walks of fixed length originates from them. The number of random walks is equivalent to the degree of the node in consideration. As there are only a few attack edges, the probability of routes from Sybils intersecting with random routes from Benign nodes is less. So, a threshold is fixed for a graph. If the number of intersections of random walks is lesser than the threshold, the suspect is a possible Sybil [10].

Drawback.

This method clearly states that there is a guarantee over the number of Sybils admitted per attack edge. Therefore, as the attack edges increases, Sybils can increase uncontrollably in the graph [1], [2].

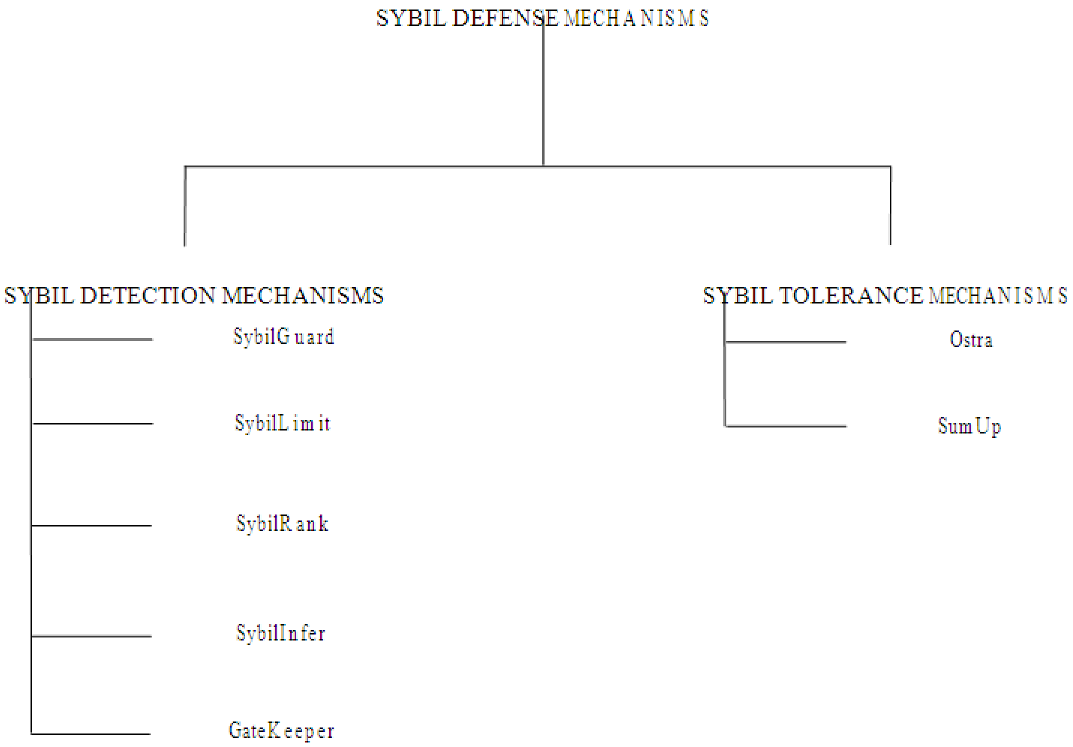

Figure 1. Topics covered in this paper

B. SybilLimit

Working.

This defense mechanism is similar to SybilGuard protocol. SybilLimit makes use of SybilGuard protocol multiple times. This protocol is that the length of random walks is smaller. By having smaller lengths, the chance of mislabeling a Sybil as Honest is lesser. Another variation in the protocol is that the last traversed edge is used to determine the intersection of random walks and not the last node [9].

Drawback.

SybilLimit works well when few attack edges join a Sybil region and an Honest region. However, with the increase in number of attack edges, more number of Sybils gain access to the Honest region and the approach becomes less efficient [1], [2].

C. SybilRank

Working.

This scheme is used in Social networks with bidirectional relationships with nodes with an intention of making manual Sybil detection easier. Honest nodes are given higher ranks. Ranks are given to nodes based on landing probability of short random walks. log (n) (early terminated power iteration) iterations are used to efficiently calculate landing probability [11]. The random walk has to start from an Honest inside

Louvain detected honest community [16]. Each community can choose its seed. Therefore, this scheme can be used in graphs with multiple Sybil and Honest communities.

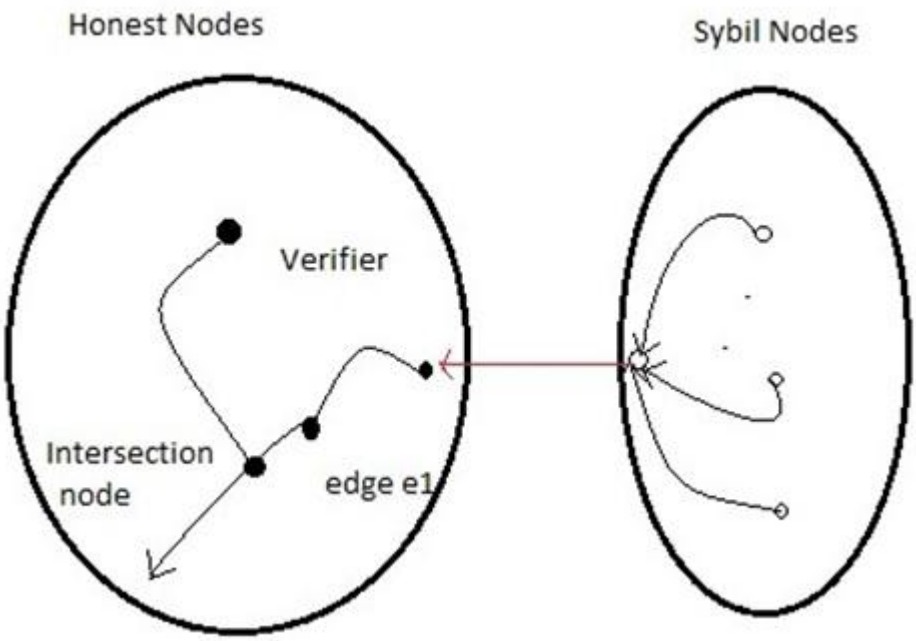

Figure 3. How a Sybil is connected to an Honest node in the Honest community. Scenario for SybilGuard. [10]

Drawback.

Although this approach is independent of the Sybil community topology, it predominantly depends on the number of attack edges. With increase in number of edges, as the number of random walks reaching Sybils also increase, Sybils would be given high ranks. As a consequence Sybils are treated as Honest nodes [1], [2].

D. SybilInfer

Working.

This protocol requires a previously determined Honest node and a Social graph as inputs. SybilInfer assumes Social graphs to be fast mixing. Research has shown that Social networks in reality are fast mixing [17]. Another assumption made is that every node knows the complete graph topology. Short, modified random walks called traces are used. Traces are represented as vertex pairs, first and last vertices. Mixing times are calculated by measuring how fast traces reach stationary distribution [13]. As discussed earlier, presence of minimum cut can be detected if the mixing time is high.

Drawback.

SybilInfer uses number of attack edges and the mixing time of the honest region to detect Sybils. Therefore, if the number of attack edges increase or if the Sybils manage disguise themselves within the Honest region, efficiency of the algorithm tends to come down [1], [2].

E. GateKeeper

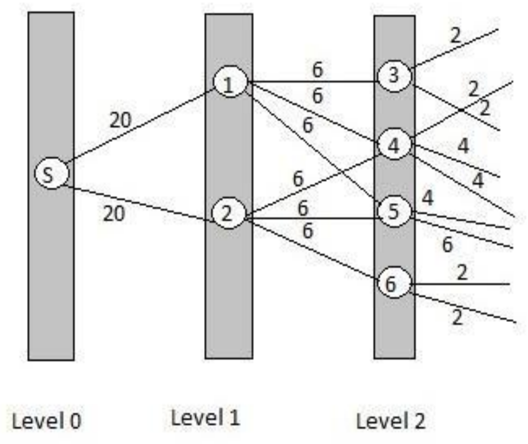

Figure 4. Ticket Distribution in GateKeeper [12].

Working.

A distributed admission control mechanism. Every node in the system that acts as an admissions controller is bound to consult with other nodes. This implies that even though a node might be a recognized admissions controller, it can be seen as a possible suspect by another node. These admissions controllers sends off certain number of tickets in Breadth First Search (BFS) manner. At every level of distribution direct neighbors receive few tickets. The tickets are discarded if a node receives tickets but has no direct neighbors. The principal assumption behind the distribution of tickets is that Sybils have access to the Honest community only through limited number of attack edges. Due the BFS distribution of tickets, nodes that are far away from the admissions controller tend to receive very few or no tickets [12].

Figure 4. shows a visualization of how the tickets are distributed in this Defense Mechanism.

Drawback.

Only if the clustering coefficient of the Sybil community is high and the access to the Honest community is limited the approach works. If attack edges increase, chances of Sybils getting would be increased as Sybils would gain more access to the admissions controller (Ticket Distributor) [1], [2].

V. SYBIL TOLERANCE APPROACHES

Sybil detection approaches are generic in nature, whereas Sybil tolerances are used for specific purposes. Instead of detecting and removing Sybils from the network these methods reduce the influence of Sybils in the network by using a credit system. Sybil tolerance approaches and corresponding drawbacks.

A. Ostra

Working.

All nodes in the network are given certain credits. Separate credits are allocated for incoming and outgoing messages. When a node receives a message, credits from all the edges in the path between the sender and receiver are reduced. The receiver has to mark the message as useful or as spam. If the message is marked as useful then the credits along the path for each edge are replenished. Spammers are blocked in this manner. Credits of edges are also replenished with time [14].

Drawback.

This method has unintended consequences, even though this method is able to block spammers. If a Sybil sends a message to another node, all the edges in the path are penalized irrespective of whether or not the intermediate edges belong to Sybils. As the system forgives spammers, Sybils are never blocked permanently. All the Sybils in the network can target and isolate certain Honest nodes by constantly spamming it [1], [2].

B. SumUp

Working.

This approach is used to detect fake votes cast in a network. Like Ostra, this approach chooses a Ticket Distributor. Here, this particular node distributes tickets and also receives votes. This node distributes tickets in a BFS manner. Each node keeps one ticket and redistributes the remaining tickets equally among the next level of BFS neighbors. Each link in the network is given a capacity equivalent to the number of tickets it has plus one. Using a feedback mechanism the vote collector is able to give negative feedback to paths from which fake votes have been cast. Too much negative feedback would lead to removal of the corresponding edge [15].

Drawback.

Adverse consequences of the feedback mechanism in this approach are similar to those of Ostra. With increase in number of attack edges, the influence of Sybils in the network tend to increase. Multiple Sybils together can work towards isolating an Honest node. Here the consequence is more severe compared to Ostra as once an edge is removed then no votes can be cast through that edge [1], [2].

VI. RESULTS

Since this is a survey paper we have no concrete section to be included as results.

VII. CONCLUSION

Since Sybil attacks are becoming widespread the need for efficient and effective Sybil Defense Mechanisms is inevitable. We have noticed from the information we have collected that all the existing Sybil Defense Mechanisms work effectively under certain assumptions about the Sybil nodes, Honest nodes and the structure of the Social Graph. But in the real world the assumptions are hardly seen to true. Therefore, existing systems for Sybil Defense are not completely infallible. In this paper we have discussed about the existing Sybil Defenses. We have also put in effort to gather information from various sources to give a concise summary of the working and drawbacks of these mechanisms.

We hope this paper could be used as reference for gaining a brief idea about the working and drawbacks of current Sybil Defenses.

ACKNOWLEDGMENT

The authors would like to express gratitude to the reviewers for their valuable time in evaluating the paper and giving important feedback, as well as Professor K. Abirami for her indispensable guidance during the course of the work.

REFERENCES

1. based Sybil Defenses", *IEEE* Trondheim, Networking Conference, 2014 IFIP, pp. 1 – 9, 2-4 June 2014.
2. H. Yu, "Sybil Defenses via Social Networks: A Tutorial and Survey," *SIGACT News*, vol. 42, no. 3, pp. 80–101, 2011.
3. B. Viswanath, A. Post, K. P. Gummadi, A. Mislove, "An Analysis of Social Network-based Sybil Defenses," in *SIGCOMM* 2010.
4. H. Gao, J. Hu, T. Huang, J. Wang, Y. Chen, "Security Issues in Online Social Networks," *IEEE* Internet Computing, vol. 15, no. 4, 2011.
5. V. Sridharan, S. Vaibhav, M. Gupta, "Twitter Games: How Successful Spammers Pick Targets," in *ACSAC* 2012.
6. Z. Yang, C. Wilson, X. Wang, T. Gao, B. Y. Zhao, Y. Dai, "Uncovering Social Network Sybils in the Wild," in *IMC* 2011.
7. L. Bilge, T. Strufe, D Balzarotti, E. Kirda, "All Your Contacts are Belong to us: Automated Identity Theft Attacks on Social Networks," in *WWW* 2009.
8. D. Irani, M. Balduzzi, D Balzarotti, E. Kirda, C. Pu, "Reverse Social Engineering Attacks in Online Social Networks," in *DIMVA* 2011.
9. H. Yu, P. B. Gibbons, M. Kaminsky, F. Xiao "SybilLimit: A Near-optimal Social Network Defense against Sybil Attacks," *IEEE/ACM* Trans. Netw., vol. 18, no. 3, pp. 885–898, 2010.
10. H. Yu, M. Kaminsky, P. B. Gibbons, A. Flaxman, "SybilGuard: Defending Against Sybil Attacks via Social Networks," *IEEE/ACM* Trans. Netw., vol. 16, no. 3, pp. 576–589, 2008.

11. Q. Cao, M. Sirivianos, X. Yang, T. Pregueiro "Aiding the Detection of Fake Accounts in Large Scale Social Online Services," in *NSDI* 2012.
12. D. N. Tran, J. Li, L. Subramanian, S.S. M. Chow, "Optimal Sybil-resilient Node Admission Control," in *INFOCOM* 2011.
13. G. Danezis and P. Mittal, "SybilInfer: Detecting Sybil Nodes using Social Networks," in *NDSS* 2009. [14] A. Mislove, A. Post, P. Bruschel, K. P. Gummadi, "Ostra: Leveraging Trust to Thwart Unwanted Communication," in *NSDI*, 2008.
14. D. N. Tran, B. Min, J. Li, L. Subramanian, "Sybil-resilient Online Content Voting," in *NSDI* 2009.
15. V. D. Blondel, J.-L. Guillaume, R. Lambiotte, E. Lefebvre, "Fast Unfolding of Communities in Large Networks," Journal of Statistical Mechanics: Theory and Experiment, vol. 2008, no. 10, 2008. unstructured networks", in N. Borisov and P. Golle, editors, Proceedings of the Seventh Workshop on *Privacy Enhancing Technologies* (*PET* 2007), Ottawa, Canada, June 2007. Springer.