Volume 4, Issue 6, June 2014 ISSN: 2277 128X

International Journal of Advanced Research in Computer Science and Software Engineering

Research Paper

Available online at: www.ijarcsse.com

Sybil-Decline: A Survey on Novel Trusted Identity and Threshold Based Path Rank for Sybil Attack Identification in Social Network

Deepti Sharma*
Department of CSE, LKCT,
Indore (M.P), India

Dr. Sanjay Thakur
Department of CSE, LKCT,
Indore (M.P), India

***Abstract*— During the last few years community network is getting popularity as per their user growth. They are generating a huge amount of data thus making promising promotion directions of business firms. Such concerns raise some issues related to fake profiles for more promotional activities. These fake profiles of the users are known as Sybil profiles and come under the Sybil attack categories. Sybil's had your personal information by using your professional and other private links. There are so many approaches developed over the last few years to overcome such issues like: SybilDefender, SybilLimit, SybilGuard, SybilShield etc. But some of the problems related to trust calculations and accurate analysis through continuous monitoring remains unaddressed. Identification and removal of such attack is not prominent and has to be done effectively.**

This work proposes a novel Path Rank, Trusted Identity, Threshold and Certificates based Sybil Detection. As the early approaches is not been able to trace the friends profile this work added an additional feature of forwarding trust and continuous behaviour monitoring. The concern is to categories the information for network partition for trusted authentic identities and unknown fake profile elements. By reasoning on these two networks, the device is then able to determine whether an unknown individual is carrying out a Sybil attack or not. At the initial level of study and problem formulation the primary solution seems to provide effective results in near future which can be evaluated practically by measuring the robustness and overhead for the system.

Keywords*— Sybil Attack, Community Network, Path Rank, Trusted Identity, Behaviour Analysis, Networked Resources, Certificates, Sybil-Decline

I. INTRODUCTION

With the tremendous growth in computing, flows of information is also getting various constraints and issues related to the modifications of authentication mechanism for user's verifications. Abrupt increase in number of nodes causes the problem of cooperative computing that requires the trust of one entity on other. In such systems where the probability of change is very high, service and user monitoring is becoming the key process for effective and secure communication. Providers services ranges is very high and the change of users identity and profiled behaviour is also non prominent. It requires high trust values between the team members for more secure relationship and communication and hence the user cooperation is also increased by using this. Transmission of data requires transferring the packets through various un-trusted nodes and devices which later on might causes data loss. If attackers control one or more participating nodes, they could modify the local raw data, local computed results, or all of the transmitted data. Clearly, by such an attacking mechanism, the attackers can modify the overall computation results of a peer-to-peer system, or even subvert the whole system. Therefore, security is a very important aspect of the research in such systems.

Distributed peer-to-peer system is the combination of various nodes demanding the information regularly. Such a heavy information flows involves the exchanges of data between different users which might not be trusted. They all are based on a common assumption that each participating user or node can controls only one identity. If the system is rejecting your identity managements then your system is affected by Sybil attack occurs mainly on distributed environment and causes identity theft or fake profile use for service acquisition. It affects the process of assumption satisfaction and evolves the generation of fake or bogus identities. These bogus Sybil identities compromise the running of the system or pollute the system with incorrect information. Most of the time these Sybil identities replaces the honest identities in a variety of tasks, including online content ranking, DHT routing, file sharing, reputation systems, and Byzantine failure defences [1].

There are some similar attacks in ad-hoc and sensor networks. Sybil identities are controlled by the adversary; she can maliciously introduce a considerable number of false opinions into the system, and subvert it, by making decisions benefiting her. Essentially, Sybil attacks break and manipulate the trust mechanism behind peer-to-peer systems. So many other works decompose existing Sybil defence schemes and demonstrate that at their core, the various algorithms work by implicitly ranking nodes based on how well the nodes are connected to a trusted node [2]. Nodes that have better connectivity to the trusted node are ranked higher and are deemed to be more trustworthy. Similarly some of the researchers had described the efforts to detect, characterize and understand Sybil account activity in the Renren online social network (OSN) [3].

II. BACKGROUND

The Sybil attack can be taken as a major security threat in distributed peer-to-peer environment where the number of nodes identities varied frequently. In this a single malicious node claims more than one identity simultaneously having the intension of distorting the usage of internet services. It generates the forged identities and affects the actual operations of the services. Let us understand its working by taking the real example of any voting system. In this Sybil create the fake identities of single users many times to outvote the operation and increases counts which later on cause the affected decision and results. This forged entries and identities dominating the operation of the distribution system. To provide the solution for pre-emption of such attacks can be described in two broad categories: centralized and decentralized. In centralized detection, a monitoring authority is responsible for providing the controls of authentications to system and uses secure credentials. Accordingly, a node can only participate in the distributed system if it is registered and its credentials are provided by the centralized authority. While in decentralized environment these authentication controls is shifted from one entity to multiple nodes involved in process and identity verifications and uses trust enabled methods. Detection of false positive and negative attacks research works on Sybil defence techniques hold the most important position [4].

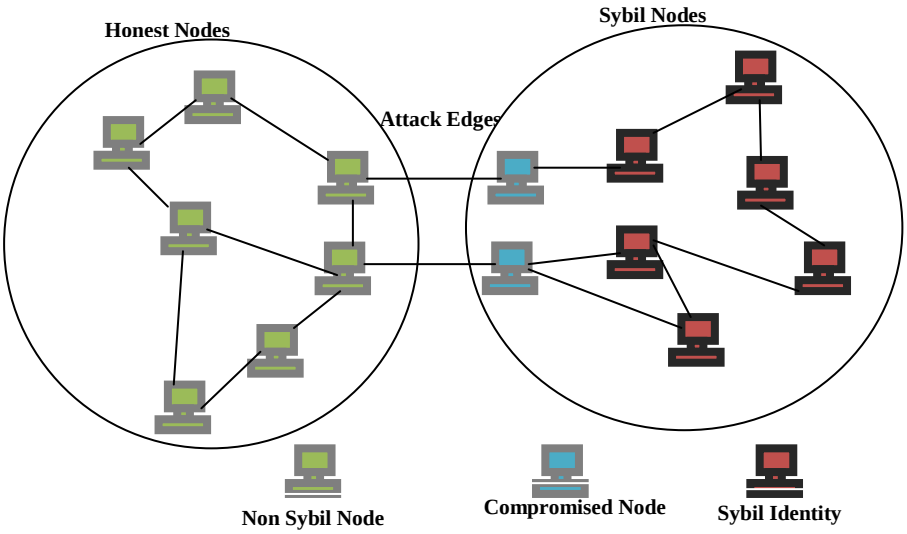

FIG 1: NODE GRAPH OF NETWORK ENVIRONMENT FOR SYBIL IDENTIFICATION

To better understand the Sybil attacks work needs to better understand the taxonomy of different types of Sybil attacks. The capability of the attacker is determined by several characteristics are given in [5]:

a) Insider vs. outsider;
b) Selfish vs. malicious;
c) Directed vs. in directed communications;
d) Simultaneously vs. gradually obtained Sybil identities;
e) Busy vs. idle;
f) Discarded or retained.

To overcome such Sybil attack occurrence in a network there are rules which needs to be reflected in execution environment by which effective and early identification of forged entries can be measured. These rules are given as:

(i) *Secure routing:* if an honest node X performs a lookup for an identifier ID, then the lookup mechanism must return the global successor of ID (present in the routing tables of honest nodes).
(ii) *Pseudonymous communication:* an adversary should not be able to determine the IP address corresponding to a user.
(iii) *Privacy of user relationships:* an adversary should not be able to infer a user's social contacts.
(iv) *Low control overhead:* the control overhead of the system should be small to enable a scalable design. This excludes flooding-based and single-hop mechanisms.
(v) *Low latency:* the length of the path used to route to an arbitrary identifier should be small, in order to minimize look up latency.
(vi) *Churn resilience:* even when a significant fraction of nodes fail simultaneously, lookup queries should still succeed.
(vii) *Fully decentralized design:* we target a fully decentralized architecture without any central points of trust/failure.

III. LITERATURE SURVEY

During the last few years various researcher had worked on identifying these Sybil attacks in timely manner. Removal of such attacks will also improve the network performance and dependencies. Out of those the major contributions are given as follows:

In 2006, Haifeng et.al suggested a new Sybil detection tool named as SybilGuard [6] and is used for minimizing the corruptive influences affects. The proposed protocol is based on the community network for communications among the

user's identities. Here the edges between the various identities is been represented by trust relationships which might be created by malicious nodes or users. Thus some mechanism needs to be designed to sole the above mentioned issues. Thus SybilGuard makes an imaginary line between the actual user and maliciously behaving users. It enables a relationship that an honest node can accepts, and also is accepted by, most other honest nodes.

In 2011, Zhuhua et.al presents a new statistical model and connected learning algorithms for identifying Sybil attacks in a collaborative network, called the latent community (LC) model [7]. In this model the whole neighborhood is partitioned in sub-networks to detect Sybil attacks. In uses a Bayesian inference approach for informing the LC model, as well as associated MCMC algorithms is also used. Evaluation of work shows experimentally that LC-based Sybil detector competes well with algorithms for the Sybil detection from the network security literature.

During the detection mechanism working for identification of Sybil attacks distributed environment consideration is must. Continuing the above process the author in [8] gives some control criteria's for early contaminations detections. The paper users trust relationship model using distributed protocol to allow nodes to collectively verify the computational work of other nodes. It also gives mechanisms to prevent the malicious influence of misbehaving nodes that do not perform the computational work. The work also proves strong evidence that Sybil- Control can be practically deployed.

The calculation of vulnerable systems for Sybils in peer to peer systems some modification is done by tool SybilLimit [9]. It leverages the same insight as SybilGuard, offers dramatically improved and near-optimal guarantees. SybilLimit's improvement derives from the combination of multiple novel techniques:

1) Leveraging multiple independent instances of the random route protocol to perform many short random routes;
2) Exploiting intersections on edges instead of nodes;
3) Using the novel balance condition to deal with escaping tails of the verifier; and
4) Using the novel benchmarking technique to safely estimate.

Finally, the results on real-world social networks confirmed their fast-mixing property and, thus, validated the fundamental assumption behind SybilLimit's approach.

SybilShield [11] is the first protocol that defends against Sybil attack utilizing multi-community social network structure in real world. It uses agent nodes for reducing the false positive rate in multiple communities by inspecting the identities of nodes. This agents helps in validation by performing random routes denies accepting the suspect node. Through the theoretical probability analysis and experiments on the MySpace data set, SybilShield is shown to greatly outperform existing mechanisms for reducing the false positive rate while keeping the effectiveness of identifying Sybil nodes with acceptable tradeoffs.

Some of the authors had also focused on data analysis based approaches for these detections such as the approach given in Sybil Identification Using KD-Tree [12]. It uses two novel algorithms SICT (Sybil identification using connectivity threshold) algorithm and SICTF (Sybil identification using connectivity threshold and frequency of visit or hitting the neighbors) algorithm. Both the algorithms are combined with previous Improved KD-Tree algorithm for community mining. Experimental comparison shows that for identifying Sybil through hitting event, the false positive is reduced than the existing random walk algorithm.

There are some other approaches which outperforms the various situations and outcomes in case of early Sybil detections before any data losses. These techniques are Privilege Attenuation [13], PrivacyJudge [14] and Trust in peer-to-peer systems [15]. Although there have been extensive research work about protecting trust schemes from various types of attacks, how to accurately and efficiently detect different attacks, for example, fake transactions, in terms of collusion remains an open problem.

In 2013, Peng et.al suggests a social network based anonymous communication in Tor and named as STor [16]. It selects the trusted router for protected transmission and anonymity. It gives the design of an input independent fuzzy model to determine trust relationships between friends based on qualitative and quantitative social attributes. STor users thus select routers by taking into account their trust in those routers. An algorithm to propagate indirect trust to a friend's friends which form a friendship circle is also given in the article. The experimental results show that STor can effectively establish trust-based circuits in a decentralized environment.

While several attacks and attack scenarios have been documented, few studies have measured the actual deployment of such attacks and none of the documented countermeasures have been tested for compatibility with an already deployed network [17].

The paper shows that the world-wide deployed KAD network suffers large number of disbelieving insertions around shared contents and enumerate them. It detects the attack after analysing the sharing of the peers' ID found around an entry after a DHT lookup. The work also evaluates the resolution and show that it detects the most efficient configurations. It is able to detect the false-negative rate, and that the countermeasures successfully filter almost all the apprehensive peers.

IV. PROBLEM STATEMENT

One In the past few years, online social networks have gained great popularity and are among the most frequently visited sites on the Web. The large sizes of these networks require that any scheme aiming to defend against Sybil attacks in online social networks should be efficient and scalable. Some previous schemes can achieve good performance on a very small network sample but their algorithms are computationally intensive and cannot scale to networks with millions of nodes.

SybilDefender [1], a centralized Sybil defence mechanism which consists of two algorithms:

- Sybil identification algorithm to identify the Sybil nodes
- Sybil community detection algorithm to detect the Sybil community surrounding

In this a Sybil node can be detected by using the above two approaches which limits the number of attack edges in online social networks. The given scheme is based on the observation that a Sybil node must go through a small cut in the social graph to reach the honest region. An honest node, on the contrary, is not restricted. Now if it starts from a Sybil node to do random walks, the random walks tend to stay within the Sybil region. Based on performing a limited number of random walks within the social graphs, the Sybil identification algorithm and the Sybil community detection algorithm are efficient and scalable to large social networks.

After studying the above mentioned approaches regarding the detection and removal of Sybil attacks or bogus identities here are the few problems which limits the capabilities of social networks are:

1) To study the Sybil Attack and its effects on social networking.
2) Evaluating the effectiveness of existing Sybil detection and tolerance algorithms
3) To develop effective Sybil detection model for identifying the adversary capabilities from a Sybil Befriend
4) To remove the limitations of relying on communities for finding Sybil's.
5) Existing decentralized defenses have largely been designed for peer-to-peer networks but not for mobile networks.

Thus for Sybil defence schemes to work well, all non-Sybil nodes need to form a single community that is distinguishable from the group of Sybil nodes. So for improved and timely identification of such attacks fake user entry needs to be identified in accurate manner. Thus this work proposes a novel model for such detection.

V. PROPOSED SYBIL-DECLINE APPROACH

According to the above identified problem area of correct prediction about the Sybil detection is to get the accurate behaviour of the social node which can be measured by difference between the fake and actual identity. If the node is having more than one identities and others than one are representing as false, than this needs to be identified before making loss to data or user profile elements. The approach is a combination of multiple steps from data generation to Sybil identifies removal. Initially, the suggested mechanism generates data regarding the profile and historical activities of the node. On the basis of this valued data individuals original entities are identified which is older and loyal with other nodes. These nodes categorized as trusted nodes which analyses the random path up to which the spreading of bogus entities is typical. Later on these random paths are calculated by assigning trusted certificates for each successful transmission or activity. It also checks whether the given trust values and path ranking is greater than a decided threshold. If it is above then it proceeds further to communicate with its more friends or else communication is aborted. The prime concern about such user and community management is for large networks which stores tera bytes of data and which has to be processed frequently to get the intelligent business decisions. Community networks spreads the information very fast so as to identify the bogus or fake profiles is very complicated task.

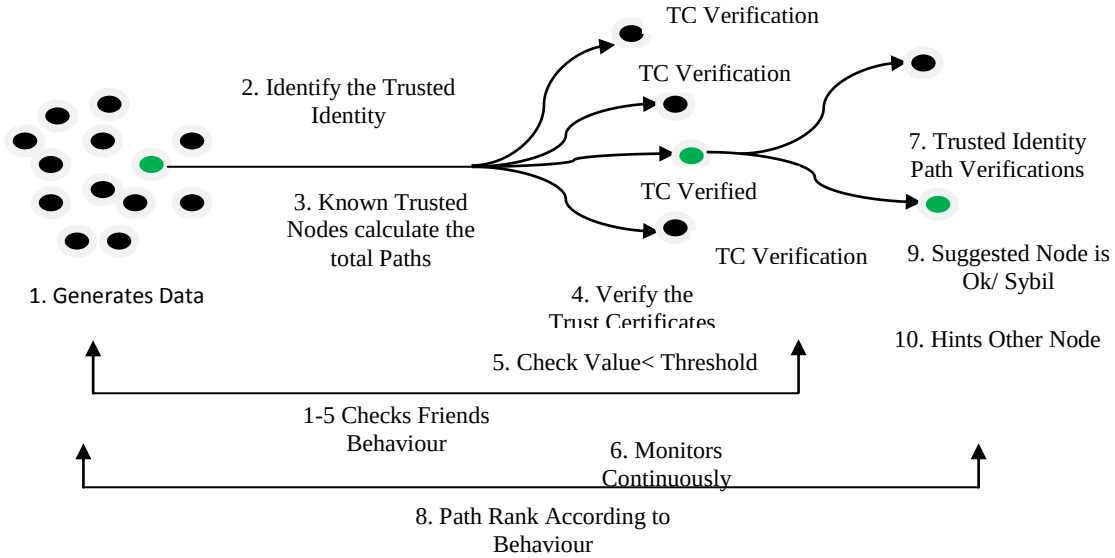

FIG 2: PROPOSED APPROACH FOR SYBIL ATTACK DETECTION & REMOVAL IN COMMUNITY NETWORK

This can be done only by confirming through some friend node. This node is called trusted behaviour identifier or trusted identity. From this identity path rank, behaviour of other users, certificates and thresholds can be easily identifiable from which accurate decision can be takes.

The idea is that a device manages two small networks in which it stores information about the devices it meets: its network of trusted friends contains trusted devices, and its network of foes contains Malicious Devices. By reasoning on

these two networks, the device is then able to determine whether an unknown individual is carrying out a Sybil attack or not. The work will also evaluate the extent to which the proposed approach reduces the number of interactions with Sybil attackers and consequently enables collaborative applications. The wok will achieve this using real mobility and social network data. It also assesses computational and communication costs on mobile phones.

Components:

(i) **Path Rank Identification:** This component ranks the different path used for communication or data transmission. The most successful path will ranked higher than the path which mostly drops the data or delay the transmissions. For this ranking some friend nodes is taken as base nodes on comparison to which the current behaviour of other nodes can be analyzed quantitatively.

(ii) **Trust Certificate Verification:** After successful communications each nodes assigns a trust certificates to its neighbour nodes. This certificate is compared with the actual ones for modification detections. If the certificates format and counts does not match with the original data or control node then it can be taken as malicious node.

(iii) **Threshold Condition Checks:** Each node and path transfers the data whose identity match is above a certain defined threshold limit.

It the limit is outreached than the authenticity of node is affected and needs to be removed. It also give the actual count of similar identities from which fake identities is deleted or preempts from its working.

(iv) **Regular Behaviour Monitoring:** The above process is regularly applied to the overall communication system. During this monitoring various local checks is been used to measure the difference and deviation from the actual behaviors.

After analysis of such Path Rank, Trust Certificates Verification, Threshold Condition and Regular Behaviour Monitoring Sybil node detection and removal can be accurately done. At the initial level of our work it seems to be giving better results than any other existing approaches. It can be implemented for so many domains like, VANET, MANET, and Social Networks etc. Later versions of this approach will definitely give improvements in social media over mobile devices where verification of entities is quite difficult.

Expected Outcomes

1. It is Helpful to find Sybil Attacks thus making the system robust.
2. Easy detection using Path rank, Trust Certificate and Threshold values.
3. Trust certificates allows only authentic user to further forward the request.
4. Business intelligence applications more likely to be motivated after Sybil removal.
5. It is used to find fake user identities.
6. It is feasible to limit the number of attack edges in online social networks by relationship rating.

VI. APPLICATIONS & EVALUATIONS

Evaluation Parameters

The goal of the proposed approach is to reject Sybil's and accept honest people. To ascertain the effectiveness of proposed work evaluation factors give correct directions:

- *Robustness:* It is measured as effective protection of community network from Sybil attacks and up to which content this detection is accurate and on time. It is also measured by false positive detection rate in unit time.
- *Overhead:* This detection must generate some of the computational complexity for the system. So these complexities must be low as per the value of information is concerned. Hence it works for time, storage, and communication overhead.

Application Domains

1. *Mobile Networks:* Observing location can differentiate between diverse devices and limits of sensible mobility can constraints attacker progress. For an attacker having a single device, all Sybil identities will always appear to move together. The defense is not applicable beyond mobile networks, and it does not protect against a single entity controlling multiple devices, each having a non-recurring cost.

2. *Auditing:* In some cases, audit can be used to find out the correctness of identity behaviour. If audit is cheap, the Sybil attack has little benefit: for instance, a large number of apparently independent identities cannot successfully convince another entity that they have factored a large number unless they have actually done so.

3. *Cash Economies: In* these, identities explicitly exchange currency for desired goods or services. In most cases, such applications are not susceptible to the Sybil attack, since they do not rely on redundancy.

4. *Reputation Systems:* For many p2p systems, including ad hoc networks and online markets, reputation systems have received a significant amount of attention as a solution for mitigating the effects of malicious peers. In an different work, various authors had evaluated the vulnerability of reputation systems to the Sybil attack, classifying them as symmetric or asymmetric approaches.

5. *Symmetric Reputation:* A symmetric reputation system is one in which an identity's reputation depends solely on the topology of the trust graph, and not the naming or identity of nodes. An attacker that wishes to increase its reputation simply uses Sybil identities to create a copy of the existing graph representing trust relationships. A symmetric reputation system cannot distinguish original nodes from the copies, and thus some Sybil node has reputations equal or better to any original node.

6. *Asymmetric Reputation Systems: In* asymmetric reputation systems, there are specifically trusted nodes from which all reputation values propagate. Alternatively, each entity separately computes a trust value along their unique paths to every other identity in the system. Since the trusted nodes cannot be impersonated, no Sybil attacker can create a duplicate graph as explained in the symmetric case. This trust value can change over time as the entity interacts with and observes the behaviour of different identities. Asymmetric reputation systems can be effective at raising the cost of Sybil attacks because attackers are forced to build up trust before effectively launching attacks. Unfortunately, these systems inevitably penalize newcomers who must prove themselves by offering benefits before getting anything in return.

VII. CONCLUSION

Abrupt increase in the user of social media and internet cause wide variation in the behaviour and frequency of content access. As this request is very large in numbers, their verification and identification of real and forge identities is moving towards critical zone. Lots of users are interested in such activities which required some fake identity creations which might be made for some disruptions to actual user's data. Thus to provide a secure mechanisms which handles such request and identities and let them secure from attackers is an emergent need of the market. Such forged identities are called as Sybil and their prevention mechanism is called Sybil attack detections. Over the last few years there are so many approaches suggested to overcome such issues of fake user behaviour but there remain some issues unaddressed. This work gives a novel Sybil attack detection and removal mechanism based on behavioral analysis, trust and threshold calculation. After prelim experiments the orientation and directions of works gives a positive outcomes and assures its applicability in near future.

VIII. FUTURE WORK

Forge identities detection is required some early decisions like verifications of user's intention at the time of profile creations. This can be achieved by historical transmission and activity details analysed in real time which demands heavy processing requirements. Thus in near future the practical implementation of suggested approach and its verifiability have to be done which later on extended in some real application.

ACKNOWLEDGEMENT

The authors would also like to thank anonymous referees for their many helpful comments, which have strengthened the paper and also like to give thank to Dr. Sanjay Thakur for discussions in specific domain.

REFERENCES

[1] Wei, Fengyuan Xu, Chiu C. Tan and Qun Li, "*SybilDefender: Defend Against Sybil Attacks in Large Social Networks*", in IEEE Transaction, Dec 2012.

[2] Bimal Viswanath and Ansley Post, "*An Analysis of Social Network-Based Sybil Defenses*", in ACM Special Issue on SIGCOMM-10, doi: 978-1-4503-0201-2/10/08, Nov 2010.

[3] Zhi Yang, Christo Wilson and Xiao Wang, "*Uncovering Social Network Sybils in the Wild*", in ACM Special Issue on IMC-11,doi: 978-1-4503-1013-0/11/11, Nov 2011.

[4] Abedelaziz Mohaisen, Huy Tran, Nicholas Hopper and Yongdae Kim, "*On the Mixing Time of Directed Social Graphs and Security Implications*", in ACM Special Issue on ASIACCS-'12,doi: 978-1-4503-0564-8/11/03, May 2012.

[5] Bimal Viswanath, Mainack Mondal, Allen Clement and Peter Druschel, "*Exploring the design space of social network-based Sybil defenses*", in IEEE Transaction, doi: 978-1-4673-0298-2/12, 2012.

[6] Haifeng Yu, Michael Kaminsky, Phillip B. Gibbons and Abraham Flaxman, "*SybilGuard: Defending against Sybil Attacks via Social Networks*", in ACM Special Issue on SIGCOMM-06, doi: 1595933085/06/0009, Sep 2006.

[7] Zhuhua Cai and Christopher Jermaine, "*The Latent Community Model for Detecting Sybil Attacks in Social Networks*", in ACM Special Issue on VLDB Endowment, Sep-2011.

[8] Frank Li, Prateek Mittal, Matthew Caesar and Nikita Borisov, "*SybilControl: Practical Sybil Defense with Computational Puzzles*", in ACM Special Issue on STC-12, doi: 978-1-4503-1662-0/12/10, Oct-2012.

[9] Haifeng Yu, Phillip B. Gibbons, Michael Kaminsky and Feng Xiao, "*SybilLimit: A Near-Optimal Social Network Defense against Sybil Attacks*", in IEEE/ACM Transaction, ISSN: 1063-6692, doi: 10.1109/TNET.2009.2034047, 2009.

[10] Yazan Boshmaf, "*A Quick Survey of Social Network-based Sybil Defenses*", in article at University of British Columbia, Vancouver, Canada.

[11] Lu Shi, Shucheng Yu, Wenjing Louy and Y. Thomas Hou, "*SybilShield: An Agent-Aided Social Network-Based Sybil Defense among Multiple Communities*", in Proceedings of IEEE Infocomm-13, doi: 978-1-4673-5946-7/13, 2013.

[12] Renuga Devi R and M. Hemalatha, "*Sybil Identification in Social Networks Using SICT and SICTF Algorithms with Improved KD-Tree*", in Journal of Theoretical and Applied Information Technology (JATIT), ISSN: 1992-8645, Vol. 56 No.2, Oct-2013.

[13] Philip W. L. Fong, "*Preventing Sybil Attacks by Privilege Attenuation: A Design Principle for Social Network Systems*", in IEEE Symposium on Security and Privacy, ISSN: 1081-6011/11, doi: 10.1109/SP.2011.16, 2011.

[14] Bastian Konings, David Piendl, Florian Schaub, and Michael Weber, "*PrivacyJudge: Effective Privacy Controls for Online Published Information*", in 3rd IEEE International Conference on Privacy, Security, Risk and Trust (PASSAT 2011), at: http://dx.doi.org /10.1109/ PASSAT/ SocialCom.2011.86, 2011.

[15] Bo Zhu, Sushil Jajodia and Mohan S. Kankanhalli, "*Building trust in peer-to-peer systems: a review*", in Int. J. Security and Networks, Vol. 1, Nos. 1/2, 2006. Pp 103-112.

[16] Peng Zhou, Xiapu Luo, Ang Chen, and Rocky K. C. Chang, "*STor: Social Network based Anonymous Communication in Tor*", in Hong Kong Polytechnic University, archive arXiv: 1110.5794v6, 2013.

[17] Thibault Cholez, Isabelle Chrisment , Olivier Festor and Guillaume Doyen, "*Detection and mitigation of localized attacks in a widely deployed P2P network*", in Springer Peer-to-Peer Networking and Applications, Volume 6, Issue 2, DOI 10.1007/s12083-012-0137-7,2012. pp 155-174.

