# Identification of Sybil Attack on Social Networks

Dr.J.I.Sheeba[1], V.Saranya[2], Dr.S.Pradeep Devaneyan[3]
*Assistant professor, Department of computer science and Engineering, Pondicherry Engineering College, Puducherry-14,*
*P.G student, Department of computer science and Engineering, Pondicherry Engineering College, Puducherry-14,*
*Professor, Department of Mechanical Engineering, Christ college of Engineering and Technology Puducherry-10,*

***Abstract - Social*** *networks play a vital role in daily life since it is vulnerable to many security attacks, namely Sybil attacks. The Sybil attack is an attack where a single user can create a many bogus identity to impersonate like others. The Sybil node in social networking is used for criminal activities such as stealing legitimate information about the user present in social networks it will lead to system degradation process. Since the Sybil identification algorithm does not provide a complete solution to detect the Sybil node in social networks. In order to overcome these drawbacks Sybil defender algorithm gets deployed in a proposed framework to detect the Sybil node in social networks. The Sybil defender algorithm will perform a limited number of random walk on social networks. Sybil defender is a combination of both Sybil identification algorithm and Sybil community detection algorithm. A Sybil identification algorithm is used to detect the Sybil node and Sybil community detection algorithm is used to detect Sybil community around the Sybil node in the social networks. By comparing with the existing approaches Sybil node will be effectively detected using Sybil defender algorithm. This proposed and existing works are measured in terms of evaluation metrics namely non-trustworthy rate, detection rate, packet loss, end to end delivery.*

**Keywords** - *Sybil attack, Sybil identification algorithm, Sybil defender algorithm, Social networks, Sybil community detection algorithm.*

## I. INTRODUCTION

The term social networks refer to the expression of a social relationship, certified or achieved, among individuals, families, households, villages, communities, regions, and so on. Each of them can play dual roles, acting both as a unit or node of a social network as well as a social actor. A Sybil attack is one in which a malicious node on a networks illegitimately claims to be several different nodes simultaneously [1].

The Sybil attack is well known in the context of peer-to-peer, wired, and wireless networks. A Sybil can delay all the messages by a forward lookup to an incorrect or non-existent peer. Finally, it can send false responses to the receiver. In Sybil attack, an attacker introduces itself in the networks with lots of identities, if an attacker gets large networks identities it can control a large portion of the networks [2].

The number of identities that an attacker can generate depends solely on the attacker's capabilities, which are limited by the bandwidth required for responding to concurrent requests by other peers in the system, the memory required for storing routing information of other peers corresponding to each and every generated Sybil identity, and computation resources required for serving concurrent requests without noticeable delay [3].

To illustrate how this attack works in real systems, imagine a recommender system built over a peer-to-peer overlay in such a system, the goal is to the filter information that is likely to be the interest of users based on others' recommendations. In that context, an attacker who can act as multiple users by faking multiple identities can easily out-vote legitimate users' votes on legitimate objects that are subjected to voting. This is almost guaranteed, given that the number of legitimate users who normally vote is always no more than 1% of the total number of users in any realistic recommendation system [4].

In the existing technique, it will detect the Sybil node by using Sybil identification algorithm since this algorithm does not provide complete solution for detecting Sybil node effectively. To overcome the weaknesses of existing technique, in this paper is proposed a Sybil defender which is a centralized Sybil defense mechanism. It consists of a Sybil identification algorithm to identify the Sybil nodes and a Sybil community detection algorithm to detect the Sybil community surrounding a Sybil node. By using these two algorithms, the numbers of attack edges are limited in the social networks [5].

The main contributions of this work includes: Based on performing a limited number of random walks within the social node, this proposed Sybil defender algorithm is more efficient than previous techniques.

The rest of the paper is organized as follows: Section2 describes related works, Section 3

describes the proposed framework, Section 4 describes experimental results and discussion and Section 5 describes the conclusion.

## II. RELATED WORKS

In this section, it proposes a survey about how to detect a Sybil attacks present in a social networks and the mechanisms.

Ankush et al., [6] proposed a parental control algorithm to detect a Sybil attack in peer to peer networks based on the reputation scheme since this algorithm applicable only for static networks.

Guojun Wang [5] is applied a Sybil identification algorithm to weeps out all Sybil peers by using neighbor similarity relationship.

Haifeng Yu [7] is introduced a Sybil guard approach. It will detect a Sybil peer present in social networks by generating a random path using pre-computed permutation.

H. Yu, P. Gibbons, M. Kaminsky, and F. Xiao [8] have proposed a Sybil limit. In this method a number of Sybil nodes accepted is reduced by a factor of minus (radicn).

N. Tran et al., [9] have proposed a Gatekeeper is another decentralized Sybil defense scheme that heavily relies on the assumption that the social networks are random expander. This is a strong assumption that has not been validated by previous research. This evaluation shows that Gatekeepers suffers from high false positive and negative rates and cannot effectively identify Sybil nodes on the real-world asymmetric social topologies.

G. Danezi et al., [10] have introduced Sybil Infer, a centralized Sybil defense algorithm, leverages a Bayesian inference approach that assigns a Sybil probability, indicating the degree of certainty, to each node in the networks. It achieves low false negatives at the cost of high computation overhead.

A.Kurve and G. Kesidis [11] proposed Sybil detection via distributed sparse cut in this method. It identifies attack edges and quarantine Sybil clusters. This method works well with dynamic trust graphs as nodes do not need to store any pre-computed data.

C.Hota [12] is proposed a safeguard algorithm, here some arbitrary verifiers are chosen. Each verifier verifies a group of arbitrary nodes, called as suspicious group, by finding paths to each suspicious node and the connection of paths is taken. After connection, the nodes remaining are more likely to be Sybil.

G. Kesidis [13] has proposed a Sybil-proof referral system, which is based on multiplicative reputation chains. Using a multiplicative reputation chain, single step and multi-step referrals can made Sybil proof.

Douceur [14] has proven that without the use of central authority, it is not possible for a system to fully defend against Sybil attack. Hence, in the p2p network, which is fully distributed, Sybil nodes cannot be removed completely from the networks.

Samidha et al., [15] proposed Sybil attack detection on p2p networks based on enhanced Sybil-resilient protocol.

Hengkui Wu et al., [16] proposed a bloom filters to detect a Sybil attack in a distributed system using a historical behavior.

In order to overcome the drawbacks of the above techniques, this proposed framework is used to detect the Sybil attack in social networks.

## III.PROPOSED FRAMEWORK

The main objective of the proposed work is to detect Sybil attack which presents in social networks. In the proposed framework, Sybil defender algorithm gets deployed to detect a Sybil attack. Sybil defender algorithm it will limit the number of random walk in a social networks. In Sybil defender algorithm consists of the following two algorithms such as Sybil identification algorithm, a Sybil community detection algorithm, and two supporting approaches which limiting the number of attack edges.

The main task of the Sybil identification algorithm is to determine whether a suspect node is Sybil or not in the social networks. Then, it shows how to efficiently detect the Sybil community around a Sybil node using Sybil community detection algorithm

The purpose of the Sybil community detection algorithm is simply supervising all the nodes in the social graph to find the Sybil community is impractical. Finally, both algorithms are built upon the deduction that the number of attack edges is limited. It is shown in Fig 1.It includes two tasks, namely [17]

1. Detection of attack
2. Sybil defender
   A.Sybil identification algorithm
   B.Sybil community detection algorithm

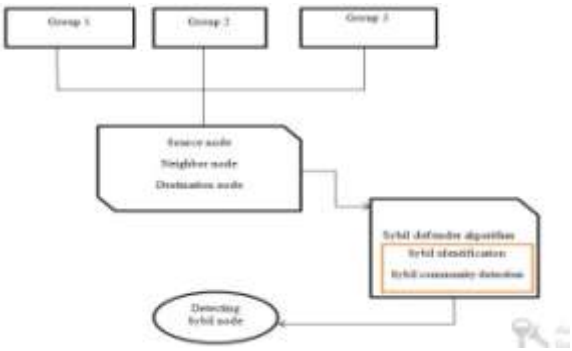

Fig 1.Proposed Framework for detecting Sybil attack in social networks

***1. Detection of attack***

In this attack is going to identify the type of attacks which is present in the networks. Naturally attacks can be classified into two types:

*Active attacks*: It will modified all the information transfer between sender and receiver

*Passive attacks*: It simply listens to all incoming and outcoming messages transfer between sender and receiver, i.e., eaves- dropping, but doesn't harm the system. A peer can be in passive mode and later in active mode [18].

***2. Sybil defender***

In the Sybil defender algorithm, it is a combination of both Sybil identification algorithm and Sybil community detection algorithm

***A.Sybil identification algorithm***

A Sybil identification algorithm that takes the social graph G (V, E), a known honest node h, and a suspect node u as input, and outputs whether u is Sybil or not. This algorithm is based on random walks. A sequence of moves of a particle between nodes of G is term as random walk. If the particle is at node i with degree $d_i$, then the probability that the particle follows the edge (i, j) and moves to a neighbor j is $1/d_i$.

The main idea behind this Sybil identification algorithm is that, as there is a small cut between the honest region and the Sybil region, the random walks originating from a Sybil node tend to get "trapped" into the Sybil region. Also, because it assumes that the size of the Sybil region is not comparable to the size of the honest region.

The number of nodes traversed by the random walks originating from an honest node will be larger than the number of nodes traversed by the random walks originating from a Sybil node, as long as the random walks are long enough to exhibit the difference between the Sybil region and the honest region, and it performs the random walks many times. For simplicity, it defines the number of times one node being traversed by a set of random walks as the frequency of that node [18].

Fig 2 shows a Sybil identification algorithm for identifying the Sybil node in the social networks. Here step 1 to 11 will performs a preprocessing process step 12 to 26 will perform random walk on social node. After completing all process Sybil node get detected.

1. j={h}
2. For i=1 to f do
3. Perform a random walk with length ls=logn originating from h
4. J=j∪U (the ending node of the random walk)
5. End for
6. L=$l_{min}$
7. While l<= $l_{max}$ do
8. For i=j.first() to J.last() do
9. Perform R random walks with length l originating from i
10 Get ni as the no of nodes with frequency no smaller than t
11 .End for
12. Output(l,mean({ni:i€ J}),stdDeviation({ni:i€J}))
13. L=l+100
14. End while
15. L=$l_0$
16. While l <=$l_{max}$ do
17. Perform R random walks with length l originating from k
18. M=the number of nodes whose frequency is no smaller than t
19. Let the tuple corresponding to length l in the outputs of algorithm 1 be (l, mean, stdDeviation)
20. If mean –m > stdDeviation* α then
21. Output k is Sybil
22. End the algorithm
23. End if
24. L=l*2
25. End while
26. Output k is honest

Fig 2 shows a Sybil identification algorithm

***B. Sybil community detection algorithm***

After one Sybil node is identified, The Sybil community detection algorithm can be used to detect the Sybil community surrounding it. The Sybil community detection algorithm takes the social graph G (V, E) and a known Sybil node as input, and outputs the Sybil community around us. The Sybil nodes can be identified by using Sybil identification algorithm or any previous scheme. It defines a Sybil community as a subgroup of G consisting of only Sybil nodes, and there is no small cut in this sub graph.

The reason it makes this definition is that if a small cut does divide the Sybil region into two parts S1 and S2, and the known Sybil nodes is s in S1, then, from the point of view of us, the honest region and S2 are similar, because there is already a small cut between S1 and the honest region and also a small cut between S1 and S2. When there is a small cut in the Sybil region, this algorithm can detect the Sybil community s.

This algorithm based on performing partial random walks originating from s. Each partial random walk behaves the same as the simple random Walks used in the Sybil identification algorithm, except that it does not traverse the same node more than once.

Therefore, when a partial random walk reaches a node with all the neighbors traversed by itself, this partial random walk is "dead" and cannot proceed.

This property makes a partial random walk originating from a Sybil node less likely to leave the Sybil region, compared with a simple random walk, because many such walks "die" when they hit the border of the Sybil region. Similar to the Sybil identification algorithm, the intuition behind this algorithm is that the partial random walks originating from a Sybil node tend to be trapped within the Sybil region, and thus, it can detect the Sybil community by examining the nodes traversed by the partial random walks. Fig 3 shows a Sybil community detection algorithm for identifying the Sybil community around a Sybil node in the social networks.

In Sybil community detection algorithm will be used to perform a walk length between node here step 1 to step 8 perform random walk and step 9 to step 13 output of walk length estimation[18].

Fig 3 shows a Sybil community detection algorithm

Finally, Sybil attack is detected on social networks by using above two algorithms namely Sybil identification algorithm and Sybil community detection algorithm.

## IV. EXPERIMENTAL RESULTS AND DISCUSSION

***Performance Metrics***

The evaluation metrics are mainly involved to calculate the effectiveness of the performance; the performance of the proposed framework is measured in terms of the quality measures namely

1. Non-trustworthy rate
2. Detection rate
3. Packet Delivery Ratio (PDR)
4. End-To-End Delay

***Non-trustworthy rate***

Non-trustworthy rate is the ratio of the number of honest peers which are erroneously marked as a Sybil peer to the number of total honest peers.

$$\textbf{Non-trustworthy} = \frac{\textbf{No.of honest peer marked as Sybil}}{\textbf{No.of total honest peer}}$$

***Detection rate***

Detection rate is the proposition of detected Sybil/malicious peers to the total Sybil/malicious peers.

$$\textbf{Detection Rate} = \frac{\textbf{Detected Sybil peer}}{\textbf{Total Sybil/malicious peer}}$$

***Packet Delivery Ratio (PDR)***

The ratio of the number of data packets receives to the packet data send to the destination. This illustrates the level of delivered data to the destination.

$$\textbf{PDR} = \frac{\sum \textbf{Number of packets receive}}{\sum \textbf{Number of packets sends}}$$

***End-To-End Delay***

The average time it takes a data packet to reach the destination. This includes all possible delays caused by buffering during route discovery latency, queuing at the interface queue. This metric is calculated by subtracting time at which first packet was transmitted from source from time at which first data packet arrived to destination.

$$\textbf{End to end delay} = \frac{\sum \textbf{(arrive time-send time)}}{\sum \textbf{Number of connection}}$$

***Existing technique:***

In existing framework, Sybil node gets detected by using Sybil identification algorithm based on neighbor similarity relationships. In this approach group of all the nodes which has similar behavior using similarity trust relationship hence it will act like an identifier source. They can send Identifiers to others as the system regulates. If a node sends less or more, the system can be having a Sybil attack node. The performance of existing framework is measured using the above metrics [18].

The below table shows a non-trustworthy rate detection and here number of inputs will be a node which is deployed. The table also shows that when number of honest peer get marked as Sybil increases than non-trustworthy also get increases.

The results were shown in the Table I. This table shows the performance of the non-trustworthy rate detection, which is detected using Sybil defender algorithm and Sybil identification algorithm.

**TABLE I.** ***Non-trustworthy rate using different inputs***

| No of nodes | Technique used | | | |
|---|---|---|---|---|
| | Existing technique | | Proposed technique | |
| | Sybil identification algorithm | | Sybil defender algorithm | |
| | No of honest peer | Non-trustworthy rate | No of honest peer | Non-trustworthy rate |
| 20 | 12 | 0.56 | 14 | 0.4 |
| 30 | 24 | 0.69 | 26 | 0.57 |
| 40 | 31 | 0.79 | 34 | 0.68 |
| 50 | 38 | 0.68 | 42 | 0.5 |
| 60 | 52 | 0.9 | 54 | 0.87 |

The below graph Fig 4 shows a non-trustworthy rate detection, which is detected using Sybil defender and Sybil identification algorithm with a number of inputs. Here x-axis denotes number of nodes deployed and y-axis denotes non-trustworthy rate detection. By observing this graph Sybil defender produces low non-trustworthy rate detection.

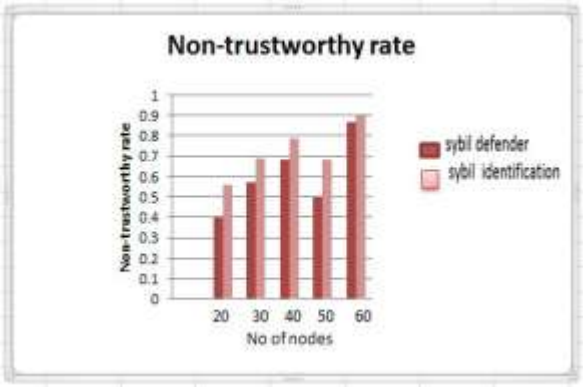


Fig 4.Graph for non-trustworthy rate using different inputs

The below table shows the performance of the detection rate, which is detected using Sybil defender algorithm and Sybil identification algorithm. The results were shown in the Table II.

**TABLE II.** ***Detection rate using different inputs***

| No of nodes | Technique used | | | | | |
|---|---|---|---|---|---|---|
| | Existing technique | | | Proposed technique | | |
| | Sybil identification algorithm | | | Sybil defender algorithm | | |
| | Total sybil peer | Undetected sybil peer | Detection rate | Total sybil peer | Undetected sybil peer | Detection rate |
| 20 | 12 | 4 | 0.5 | 10 | 9 | 0.75 |
| 30 | 15 | 8 | 0.6 | 15 | 13 | 0.6 |
| 40 | 26 | 15 | 0.63 | 30 | 10 | 0.63 |
| 50 | 35 | 21 | 0.789 | 29 | 20 | 0.789 |
| 60 | 46 | 28 | 0.85 | 38 | 28 | 0.85 |

The below graph Fig 5 shows a detection rate, which is detected using Sybil defender and Sybil identification algorithm with a number of inputs. Here x-axis denotes number of nodes deployed and y-axis denotes detection rate. By observing this graph Sybil defender produces high detection rate

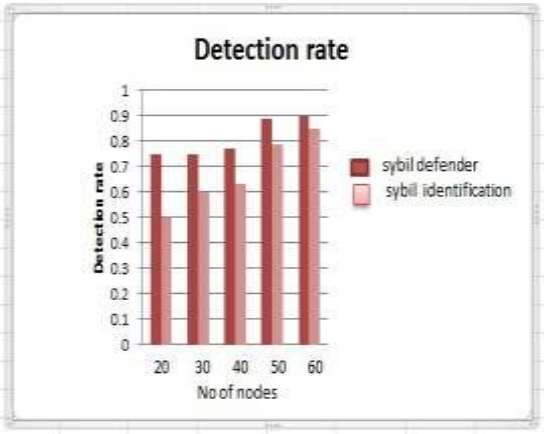


Fig 5.Graph for detection rate using different inputs

The below table shows the performance of the packet loss, which is detected using Sybil defender algorithm and Sybil identification algorithm. The results were shown in the Table III.

**TABLE III.** ***Packet loss using different inputs***

| No of nodes | Technique used | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Existing technique | | | | Proposed technique | | | |
| | Sybil identification algorithm | | | | Sybil defender algorithm | | | |
| | No of packet sends | No of packet receive | Packet delivery | Packet loss | No of packet sends | No of packet receive | Packet delivery | Packet loss |
| 20 | 20 | 15 | 0.75 | 10 | 22 | 11 | 0.7 | 7 |
| 30 | 30 | 13 | 0.6 | 20 | 35 | 30 | 0.69 | 14 |
| 40 | 40 | 40 | 0.625 | 9 | 37 | 30 | 0.64 | 5 |
| 50 | 50 | 50 | 0.61 | 14 | 55 | 50 | 0.70 | 12 |
| 60 | 60 | 60 | 0.5 | 10 | 57 | 55 | 0.89 | 8 |

The below graph Fig 6 shows packet loss, which is detected using Sybil defender and Sybil identification algorithm with a number of inputs. Here x-axis denotes number of nodes deployed and y-axis denotes packet loss. By observing this graph Sybil defender it produce low packet loss.

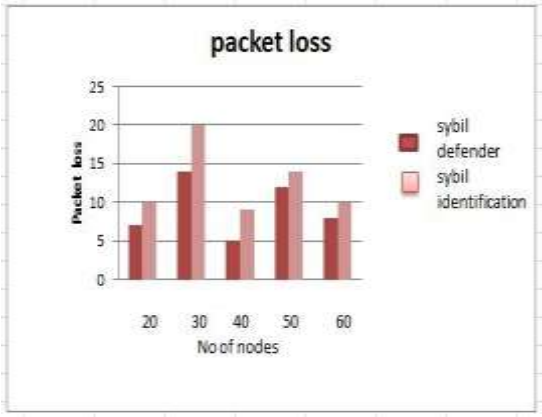


Fig 6.Graph for packet loss using different inputs

The below table shows the performance of the end to end delay, which is detected using Sybil defender algorithm and Sybil identification algorithm. The results were shown in the Table IV.

**TABLE IV.** ***End to end delay using different inputs***

| No of nodes | Technique used | | | | | |
|---|---|---|---|---|---|---|
| | Existing technique | | | Proposed technique | | |
| | Sybil identification algorithm | | | Sybil defender algorithm | | |
| | Send time | Arrive time | End to End delay | Send time | Arrive time | End to End delay |
| 20 | 10 | 15 | 0.89 | 14 | 14 | 0.7 |
| 30 | 2 | 4 | 0.8 | 9 | 9.1 | 0.78 |
| 40 | 1 | 3 | 0.79 | 4 | 4.2 | 0.67 |
| 50 | 2 | 16 | 0.6 | 7 | 7 | 0.58 |
| 60 | 5 | 6 | 0.78 | 8 | 8 | 0.64 |

The below graph Fig 7 shows end to end delay, which is detected using Sybil defender and Sybil identification algorithm with a number of inputs. Here x-axis denotes number of nodes deployed and y-axis denotes end to end delay. By observing this graph Sybil defender it produces a low end to end delay rate.

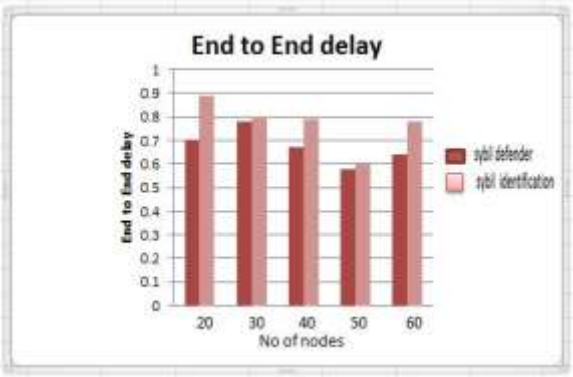


Fig 7.Graph for end to end delay using different inputs

## V CONCLUSION

Social networks are vulnerable to security attack, namely Sybil attack. Sybil attacker can create fake accounts in social networks for stealing legitimate user information it can be detected by using Sybil defender algorithm. The Sybil defender algorithm is evaluated in terms of performance metrics namely non-trustworthy rate, detection rate, packet loss and end to end delivery. By comparing with existing technique Sybil defender will be provided an effective result.

## ACKNOWLEDGMENT

I express the sincere thanks to our Institution for extending the infrastructural facilities to carry our work successful

## REFERENCES

[1] Rakesh G.V et.al "A Survey Of Techniques To Defend Against Sybil Attacks In Social Networks" in International Journal of Advanced Research in Computer and Communication Engineering Vol. 3, Issue 5, May 2014

[2] Zied Trifa "Sybil Nodes as a Mitigation Strategy against Sybil Attack" in International Workshop on Secure Peer-to-Peer Intelligent Networks & Systems (SPINS-2014).

[3] Jochen Dinger and Hannes Hartenstein, "Defending the Sybil Attack in P2P Networks: Taxonomy, Challenges, and a Proposal for Self Registration", In Proc. First International Conference on Availability, Reliability and Security (ARES 2006), Vienna, Austria, 2006, pp. 756-763.

[4] RoopaliGarg" Comparison between Sybil Attack Detection Techniques: Lightweight and Robust in International Journal of Advanced Research in Electrical, electronics and Instrumentation" Engineering Vol. 3, Issue 2, February 2014.

[5]Guojun Wang "Neighbor Similarity Trust against Sybil Attack in P2P E-Commerce"in IEEE transactions on parallel and distributed systems, vol. 26, no. 3, March 2015.

[6] Ankush Tehale "Parental control algorithm for Sybil detection in distributed p2p networks" in International Journal of Scientific and Research Publications, Volume 2, Issue 5, May 2012.

[7] Haifeng Yu, M. Kaminsky, P. B. Gibbons, A. Flaxman, "SybilGuard: Defending Against Sybil Attacks via Social Networks", in Proc. ACM SIGCOMM, 2006, pp. 267–278.

[8]Haifeng Yu, Phillip B. Gibbons, M. Kaminsky, F. Xiao, "SybilLimit: A Near-Optimal Social Network Defense against Sybil Attacks", in Proc. IEEE/ACM transactions on networking, vol. 18, no. 3, June 2010.

[9] N. Tran, J. Li, L. Subramanian, and S.S. Chow, "Optimal Sybil Resilient Node Admission Control," Proc. IEEE infocom, 2011.

[9]A. Cheng, and E. Friedman,"Sybil proof reputation mechanisms", In Proc. ACM SIGCOMM Workshop on Economics of Peer-to-Peer Systems", 2005, ACM Press, pp. 128-132.

[10] G. Danezis and P. Mit, "Sybil infer: Detecting Sybil Nodes Using Social Networks," Proc. Network and Distributed System Security Symp. (NDSS), 2009.

[11] A. Kurve and G. Kesidis, "Sybil Detection via Distributed Sparse Cut Monitoring", in Proc. ICC 2011, pp. 1-6.

[12] C. Hota, J. Lindqvist, K. Karvonen, A. Ylä-Jääski, Mohan C.K.J"Safeguarding Against Sybil Attacks via Social Networks and Multipath Routing", in Proc. NAS 2007,pp. 122 – 132.

[13] G. Kesidis, A. Tangpong and C. Griffin, "A Sybil-proof referral system based on multiplicative reputation chains," IEEE comm. letters, pp. 862- 864, nov. 2009.

[14] J. Douceur" The Sybil Attack. In 1st International Workshop on Peer-to -peer Systems" (IPTPS '02). Springer,, 2002, pp. 251-260.

[15] Samidha Nagdeve "Sybil attack detection on peer to peer network based on enchanced Sybil –resilient protocol "in international journal for scientific and research and development..
[16] Hengkui Wu" SybilBF: Defending against Sybil Attacks via Bloom Filters in ETRI Journal, Volume 33, Number 5, October 2011
[17] Wei , Fengyuan Xu SybilDefender: "A Defense Mechanism for Sybil Attacks in Large Social Networks" in IEEE transactions on parallel and distributed systems, vol. 24, no. 12, December 2013
[18] Saranya V et.al,"Detection of Sybil attack on social networks using Sybil defender algorithm" International journal of control theory and applications (In press).