Journal of Theoretical and Applied Information Technology
20th October 2013. Vol. 56 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

SYBIL IDENTIFICATION IN SOCIAL NETWORKS USING SICT AND SICTF ALGORITHMS WITH IMPROVED KD-TREE

[1]RENUGA DEVI. R, [2]M. HEMALATHA
[1 2]Department of Computer Science,
Karpagam University, Coimbatore, INDIA.
E-mail: [1]nicrdevi@gmail.com, [2]csresearchhema@gmail.com

ABSTRACT

Most of the large scale social networking sites and small private social networks on the Internet are open to Sybil attacks. Lack in powerful user identity yields these systems at risk to Sybil attacks. A large number of methods have been proposed to solve this problem, but each method differs greatly from other based on the algorithms which they used, and network. In this paper we proposed two novel algorithms to identify the Sybil nodes in network community. We proposed SICT (Sybil identification using connectivity threshold) algorithm with Improved KD-Tree. Connections between the nodes are established, and connection threshold is compared with each node, if the connection establishment is exceeding the threshold then the node is identified as Sybil. We proposed SICTF (Sybil identification using connectivity threshold and frequency of visit or hitting the neighbors) algorithm, where the maximum variance of connectivity, length and frequency of a node can be calculated for a particular time interval and the maximum variance with respect to connectivity, length, and frequency is said to be Sybil. Both the algorithms are combined with previous Improved KD-Tree algorithm for community mining. Experimental results show that proposed SICTF algorithm performs well compared to the existing algorithm.

Keywords: *Social Networks, Sybil Node, Community Mining, Improved KD-Tree, SICT, SICTF*

1. INTRODUCTION

Social networks are gaining more popularity recently. Communication between the users of networks only requires the users to be part of the same network. Because of Popularity and existence of large number of users and their communication privacy issues are raising. All kind of distributed systems are vulnerable to Sybil attacks. Open access systems like Facebook, Twitter, LinkedIn, YouTube, and Torrent allows all users on the Internet to access the system easily without any limitations. There is lack in strong user identity, so it makes the open access systems easily vulnerable to Sybil attacks [1]. An attacker can easily create a number of duplicate or fake identities (called as Sybil) to corrupt the system with fake information and affect the accurate performance of the system.

Sybil attacks are found in assorted domains, from aegis and acquisition in peer-to-peer networks to collaborative voting and advocacy systems. There are two adulatory approaches for ambidextrous with Sybil attacks. The aboriginal is prevention: architecture aegis mechanisms that accomplish it absurd for attackers to accretion admission to the arrangement in the aboriginal place, usually through character analysis schemes [2].

Now a day there has been abundant action in the analysis association over application amusing networks to abate assorted identity, or Sybil, attacks [3]. Avoiding assorted identity, or Sybil, attacks are basic problem in the architecture of broadcast systems [4]. Malicious attackers can actualize assorted identities and access the alive of systems that stay aloft accessible membership. An amount of schemes accept been proposed that attack to avert adjoin Sybil in an amusing arrangement by application backdrop of the amusing network structure [5, 6, 7, 8].

Most of the Social network Sybil protection methods are assumed that the attacker can develop a random Sybil identity in large scale social networks. And it assumed the attacker cannot create a random number of nodes of social network connections with the non Sybil nodes. The protection methods took results in assumptions like, the Sybil nodes are poorly connected with the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

remaining portion of the network but non- Sybil nodes are strongly connected with the network. A link within the social network between any users represents a confidence relationship between those users. It is affordable to assume that an attacker typically has few links to honest users since establishing faith links needs vital human efforts. Therefore, Sybil resilient admission management may be declared as follows: contemplate a social network consisting of honest users Associate in Nursing every which way several Sybil connected to honest nodes via attack edges (an attack edge could be a link between an honest and a Sybil node). Given Associate in Nursing honest node acting because the admission controller, verify the set of nodes to be admitted so the overwhelming majority of honest nodes in are admitted and few Sybil nodes are admitted [9].

In the past few years, on-line social networks have gained much popularity and are among the foremost oftentimes visited sites on the net. The big sizes of those networks need that any theme going to defend against Sybil attacks in on-line social networks ought to be economical and scalable. Some of the previous schemes can do smart performance on an awfully little network however their algorithms are computationally intensive and can't scale to networks with several nodes [10].

The existing model depends on partitioning the network into communities, those square measure subsets of the network that have robust internal connections. Community detection in graphs isn't a replacement problem; actually, it's a widely-studied and mature field with long history. We are going to cowl a number of the connected work from this space of analysis subsequently; the essential survey on the world however whereas community detection is widely-studied, no community detection methodology is clearly applicable to the matter of automatic Sybil detection. To notice Sybil, it's not enough to partition the network into tightly-connected communities; those communities should at the same time be analyzed to visualize however they connect with the remainder of the network.

In this paper we proposed sensible techniques to find attack edges in social networks, and community mining based on the filtered nodes. The survey results shows that the belief created by previous work that every relationship in social networks square measure trustworthy doesn't hold in social networks, and it's possible to find the attack edges in social networks by relationship rating. We proposed SICT algorithm with Improved KD-Tree. In this connections between the nodes are established, and connection threshold is compared with each node, if the connection establishment is exceeding the threshold then the node is identified as Sybil node. Next we proposed SICTF algorithm with Improved KD-Tree, here the maximum variance of connectivity, length and frequency of a node can be calculated for a particular time interval and the maximum variance with respect to connectivity, length, and frequency is said to be Sybil nodes.

2. LITERATURE SURVEY

Now a day there has been several analyses over social networks to decrease assorted identity, or Sybil, attacks [3]. Several numbers of methods are proposed, but they alter abundantly in the algorithms they use and in the networks aloft which they are evaluated. In 2011, Zhuhua Cai and Christopher Jermaine proposed LC model. The nodes in a Social network are clustered into several communities. The set of nodes are relatively interconnected with each other. Every community in a network is connected with a hidden position in a multi-dimensional Euclidean space, so they used the name latent community mode; the location of each community shows how the nodes are connected with each other communities in the network. Network communities which are very close to other nodes that have several links between them; the nodes which are away from each other communities would have very few links. The authors mentioned this setup as the subsequent stochastic procedure underlies the LC Model:

Step 1: For each community.

Step 2: For each community, the number of edges connecting internal nodes is generated as

Step 3: For each pair of distinct communities, the number of cross-community edges is generated as:

The above process explained as follows, through step (1) by drawing the location from a randomly generated variable which is having distribution it places all communities in the space. From Step (2), each set of nodes in the network community are paired with the probability. And in step (3), based on the distance between the community, set of nodes from various communities are linked each other, and represents the Euclidean distance between the latent positions linked with the network communities. So the probability value of two nodes inside the network communities are connected drops exponentially with increasing the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Euclidean distance between the network communities, here is simply a scaling factor.

From the analysis LC model provide better results for automatically finding the Sybil attacks in a network. Compared to other methods, LC model is the best detection method. Even though it provides better results it does have few weaknesses. That is the LC method does not work better under a tree-topology attack [2].

Nguyen Tra et al., proposed Gatekeeper method. It is a decentralized Sybil resilient admission control protocol. It considerably improves over Sybil Limit. The proposed system contains honest nodes, which is belongs to honest users. An undirected graph among all nodes in the network system are exists. The connection between the two honest nodes or users reflects the confidence relationship between the users in the real-world. The information of the social graph is scattered among all nodes. Mainly, each trust node knows their immediate neighbor nodes on the social graph and such nodes may not have the knowledge about the rest of the nodes in a graph. Each node has a public and private key which is locally generated. Every node must knows the public-keys of its neighbor node, but, there is no public-key communications that allows a node to properly find out of all other nodes public-keys. But still in the appearance of a numerous number of attack edges, Gatekeeper method can considerably limit the few number of admitted Sybil nodes per attack edge [9].

Wei Wei, et al., proposed SybilDefender algorithm to efficiently categorize the Sybil nodes and find the Sybil communities in the region of a Sybil node, yet the amount of Sybil nodes produced by every attack edge is similar to the tentatively measurable lesser bound. The authors proposed two methods to limit the amount of attack edges in Social networks. SybilDefender algorithm is a centralized method for Sybil defense method. It is a combination of three main components. Those are, a Sybil Identification method or algorithm to find the Sybil nodes, a Sybil community detection method or algorithm to find the Sybil community nearby a Sybil node, and two basic methods to limit the amount of attack edges and nodes in Social networks. These methods are based on few observations. That are, a Sybil node have got to go throughout a little cut in the social graph to reach the honest node or their region. The authors choose the Sybil node to do random walks, the random walks be likely to keep on within the Sybil region [10].

Viswanath et al., has performed an analysis study on SybilLimit, SybilGuard, SumUp and SybilInfer. Their study shows that two possible boundaries of social-network based admission control system. Existing protocols falsely refuse several honest nodes as Sybil nodes in several small networks which is having up to tens of thousands of nodes reveal community structure but not fast-mixing. Form their analysis they suggests that Sybil resilient admission control method should performed only on large-scale social networks. If the larger the graph, then it provides the better connected communities to each other and mixing time also faster than other [11].

G. Danezis and P. Mit proposed SybilInfer algorithm, is a centralized Sybil protection algorithm, it is a Bayesian inference method that calculate a Sybil probability value by representing the degree of conviction to every node in the network. It gains low false negatives value at the elevated computation overhead. The total time complexity of SybilInfer method is denoted as , where is the collection of nodes in the social graph. SybilInfer method handled very small social networks up to 30 thousand nodes [12]. Xu et al., proposed an algorithm to calculate the shortest path between the pair of nodes within the given network in every round, which makes it unreasonable for yet small sized social networks. In difference, SybilDefender method only works on performing a few numbers of arbitrary walks in the social network graph, also this method is scalable to large scale networks [13].

Several existing Sybil defense methods include a plenty of useful and sensible optimizations that improve the system performance in particular application scenarios. Best examples are, SybilGuard method [14] and SybilLimit [15] have a plenty of design facilities that make easy of their uses in decentralized network systems. Likewise, SumUp method [16] has optimizations techniques to online comfortable voting systems. On the other hand, main goal of the Sybil protection method is to discover the center graph partitioning algorithm. Usually, open access systems supported a elementary authority like CAPTCHA or process puzzles to mitigate the Sybil attack [17], [18], [19]. But sadly, these solutions can solely limit the speed with that the assailant will introduce Sybil identities into the system rather than the full range of such identities. Even before the recent surge of interest in social-network-based Sybil defenses, there are makes an attempt at exploiting the trust graph among users to mitigate the Sybil attack: Advogato

Journal of Theoretical and Applied Information Technology
20th October 2013. Vol. 56 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

methods [20], Appleseed methods [21] and SybilProof methods [22] square measure the most well-known of those early proposals. However, it is not the goal of those protocols to perform Sybil-resilient node admission. Rather, they aim to calculate the name of every user/node during a manner that stops the assailant from boosting its name victimization Sybil identities.

Community Detection issues, in contrast to Sybil defense, has been a well and long time studied topic in social science, biology, arithmetic, and physics. Many researches are going on this community mining downside. Such as, Fortunato's current survey summaries many approaches for community detection with numerous ways that to live the quality of communities [23]. In Jure Leskovec and his colleagues by experimentation compare a variety of community detection ways supported many common objective functions [24]. Xu et al., proposed an algorithm to calculate the shortest path of the nodes between every set of nodes within the social network in every round. But it is unreasonable for yet small sized social networks. In difference, SybilDefender methods based on the stage of performing a very few number of random walks in the social graph, Also it is most scalable methods to large scale networks [25].

3. IMPROVED KD-TREE ALGORITHM

An improved KD-tree with LM algorithm was proposed early with the joint encoding scheme to reduce the memory limitations. And the stopping criterion is calculated automatically by efficiently determining the minimum Eigen-gap without explicitly computing eigenvalues. Improved KD-Tree based clustering method uses the previous standard quad tree coding method followed by the neighbor joint coding algorithm. It selects neighbors must be coded jointly or else separately. Once the algorithm reaches the leaf information then the algorithm seems for the neighbor leaf, which is previously transferred or selected by the algorithm. Once the algorithm finds a shifted neighbor then it calculate the sample parameters for all the leaves in terms of a perfect assessment metric. One possible alternative of evaluation metric might be a distance - error. The distance-error among the sample parameters of the leaf and its neighbor is at intervals some predefined space metric, then the incoming leaf data won't be transmitted and also the neighbor joint writing variable are going to be set to at least one pursued by the two bits for the neighbor index (considering solely k neighbors), otherwise leaf data are going to be transmitted and also the neighbor joint writing variable are going to be set to zero. These neighborhood sets can currently be distended to incorporate the closest neighbors. The below Table 1 shows the Improved KD-Tree algorithm and it process.

Table 1: Improved KD-Tree Algorithm

Step 1: Create an Improve KD-Tree for the given data $x_i, i = 1,.....n.$

Step 2: For $j = 1,.....q.$ calculate the rank density p_j of each leaf bucket L_j. // Where q is the leaf buckets in KD-Tree.

Step 3: Calculate mean value, m_j

Step 4: Choose

$c_1 = m_z, where\ z = \arg_j \max\ p_j$

Step 5: For $t = 2,.....k,$ and for $j = 1,.....q$ evaluate

$g_j = \{\min_k = 1,.....t[d(c_k, m_j]), P_j\}$

$c_t = m_2\ where\ z = \arg_j \max(g_j).$

Step 6: go to step 3, until the convergence criteria is met and calculate a second possible list of K initial centers $(c_1, c_2, ...c_K).$

Step 7: Return $(c_1, c_2, ...c_K).$

3.1 Clustering Results Using Improved KD-Tree with LM Algorithm

The dividing planes next to any path starting the root to a different node illustrate an exclusive box-shaped section of space, and each following plane divide this box into two different boxes. Each box-shaped area is defined by k planes; here k is the amount of dimensions. In normal KD-Tree algorithm, it split the plane in to two dimensions based on the 2k. An Improved KD-Tree algorithm the K values are determined by us, so the clustering process much faster than the traditional KD-Tree algorithm.

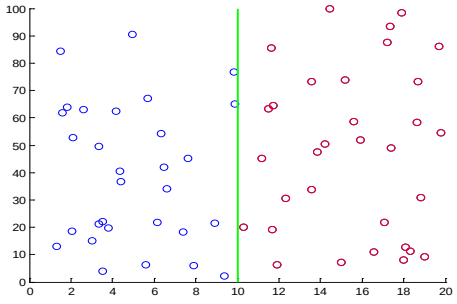

Figure 1: Clustering Result Using Improved KD-Tree

Journal of Theoretical and Applied Information Technology
20th October 2013. Vol. 56 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4. SYBIL IDENTIFICATION USING SICT AND SICTF ALGORITHMS WITH IMPROVED KD-TREE

Existing SybilDefender algorithms are fully based on the frequency of nodes. So we proposed a new algorithm based on the hitting event instead of frequency of nodes. The proposed Sybil Identification algorithms SICT and SICTF can detect Sybil attacks by observing hitting event distribution of a suspect node over a period of time intervals. It randomly selects nodes from the originator node with length. And observe the hitting event of each node with its neighbors in the periodical time. The neighboring nodes of observed nodes within their length also observed for the hitting event of with their neighbor lists. The repeatedly select the random nodes until the stopping criterion is met. The measurement of average hitting event all time interval is calculated for each node based on hitting count with its neighbors. After the measurement is over, the collected measurements from all nodes, the node can locally compute an estimated hitting event of all neighbor nodes, if hitting event is exceeds a predefined threshold for any nodes, that node is considered as Sybil node. If identifying Sybil through hitting event, the false positive is reduced than the existing random walk algorithm.

4.1 SICT Algorithm with Improved KD-Tree (Sybil Identification Algorithm Using Connectivity Threshold)

The proposed Sybil Identification algorithm using Threshold value, analyzes the nodes in the given network, then we established the connection Cn for each nodes in the network. Every network community consists of honest h nodes as well as Sybil or attacker node s1. The connectivity established between each and every nodes counted in a frequent time interval. The connection threshold is compared with the connection count of each node. If the connection establishment is exceeding the threshold then the node is identified as Sybil node. The proposed algorithm is given below.

4.1.1 SICT with Improved KD-Tree algorithm (n,Req, Res, xi,j)

Step1: Initialize connection threshold C_{th},

Set $t_i = 0$

Step2: Check the node is registered or not

Do

$t_i = t_{i-1} + \Delta t$ where i>0 //changing time interval

Step3: If registered (n)

Access(n) //Access permitted

Else

Denied (n) // Denied permission

Step 4: Connection established via request and response of a node.

$Req_i \rightarrow j$ // node I requesting node j

$Res_j \rightarrow i$//node j responded to node i

Step 5: If $Req_{i\ AND}\ Res_j$ is true

$x_{i,j} = x_{i,j} + 1$ // connected node

Else

$x_{i,j} = 0$// connection less node

Step 6: Repeat step 2 to step4

Step7: $y_{i\ =1,2,3..n} = \sum_{j=1}^{n} xij$

Step 8: if $y_i > C_{th}$ Then y_i is identified as Sybil

Step 9: End

Journal of Theoretical and Applied Information Technology
20th October 2013. Vol. 56 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4.2 SICTF Algorithm with Improved KD-Tree (Sybil Identification Algorithm Using Connectivity Threshold and Frequency of Visit or Hitting the Neighbors)

The proposed Sybil Identification using connectivity threshold and frequency of visit or hit the neighbor node algorithm works on the following manner. It considers the established connectivity Cn and we Initially Set the time ti where ti =0, by changing the time interval ti = ti-1 + where i>0 we can calculate connectivity at a particular time interval ti.. Then we calculate the length l and frequency f of a node at a particular time interval, by frequently changing the time interval we have to calculate the connectivity. The maximum variance of connectivity, length and frequency of a node can be calculated for a particular time interval. The change in variance of the connectivity, length and frequency for respective time interval can be noted. The maximum variance with respect to connectivity, length, and frequency is said to be Sybil nodes. The proposed algorithm is given below.

4.2.1 SICTF with Improved KD-Tree Algorithm

(Cn, ti, x)

Step1: Consider the established connection C_n

Step2: Set $t_i = 0$ // Initial time

Step 3: Do $t_i = t_{i-1} + \Delta t$ where i>0

//changing time interval

Calculate $C_{n(i)}$// follow algorithm 1

Calculate $l\ (C_{n(i)})l(C_{n(i)}) = \frac{\sum \forall (x_{i,j}=1)}{Cn(i))}$

//Length of a established connection in a given time interval

Calculate f ($C_{n(i)}$)

f(n') = $\sum v(n_i)$ //frequency of node

f(n)→ $\sum f(n')$

f ($C_{n(i)}$) = $\frac{f(n)}{Cn(i))}$

While(t_i = x) // x is target time

Step4: If

$t_i \rightarrow \forall$ (max(($C_{n(i)}$) ,l ($C_{n(i)}$) , f ($C_{n(i)}$))) then

$S_i \leftarrow \forall$ (max (($C_{n(i)}$) ,l ($C_{n(i)}$) , f ($C_{n(i)}$)))

// Connection made by Sybil is found.

Step 5: End

In following Table 2, we are given the notation list which is used in the algorithms.

Table 2: Notations Used in the Algorithms

Cn	Global connection
n	Node
Access(n)	Access permitted to node
Denied (n)	Node Denied
Reqi	Requesting Service
Resj	Responding service to the node
xi,j	Connection Between Two Nodes
ti	Time Period
$C_n(i)$	Established connection in a particular time interval
l ($C_n(i)$)	Length of established connection in a particular time interval
v(ni)	Visited Node
f(n')	Frequency Of a Node
f(n)	Frequency Of all Nodes
f ($C_n(i)$)	Frequency of established connection in a particular time interval
Si	Sybil Connection

5. EXPERIMENTAL RESULTS AND DISCUSSIONS

We have taken Two Input datasets for evaluation. Those are dolphin and Wiki vote data sets. For this paper we have taken dolphin data set to evaluate the algorithm performance. The dolphins.gml file posses an undirected social network of recurrent links among 62 dolphins in a neighborhood living off suspicious Sound at New Zealand. Lusseau et al. (2003) compiled these dataset for their use. We have taken this dolphin dataset for this paper. The clustering results after eliminating Sybil nodes are compared with our previous work Improved KD-Tree method based clustering based on few parameters. The proposed algorithm is examined by the basic parameters that are accuracy, recall, precision, Roc curve.

An accuracy value is represented as the degree of

Journal of Theoretical and Applied Information Technology
20th October 2013. Vol. 56 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

closeness of dimensions of a capacity to that quantity's true value. The precision values of a system are also called repeatability or reproducibility. Precision value is the degree to which frequent measurements below unaffected circumstances shows the similar results. We can calculate the precision value using the $\Pr ecision = \frac{TP}{TP + FP}$ formulae. Recall value is considered as based on the recovery of information at true positive calculation, false negative. We can calculate the recall value using the $\mathrm{Re}call = \frac{TP}{TP + FN}$ formulae. Receiver Operating Characteristic, called as ROC curve. ROC is a graphical diagram. It used to demonstrate the performance of a system.

The evaluation results of the SICT and SICTF algorithms are given in the following figures. Both the proposed algorithms are combined with the Improved KD-Tree algorithm for better community mining process

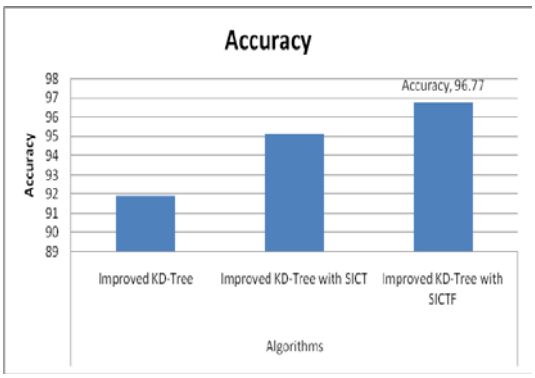

Figure 2: Accuracy Comparison

Figure 2 shows that the comparison of accuracy values of the system between the Improved KD-Tree algorithm and the Sybil Identification Algorithms. Accuracy values represented at Y-axis in percentage values (%), and the algorithms are represented in X-axis. The previous Improved KD-Tree algorithm gives 91.94% of accuracy. The proposed SICT algorithm with Improved KD-Tree provides 95.16%. But the SICTF algorithm with Improved KD-Tree gives 96.77% accuracy results. The accuracy value of the proposed SICTF with Improved KD-Tree algorithm provides higher accuracy than the Improved KD-Tree algorithm. Finally our proposed algorithm achieves a high level of the accuracy value compared to the previous algorithm.

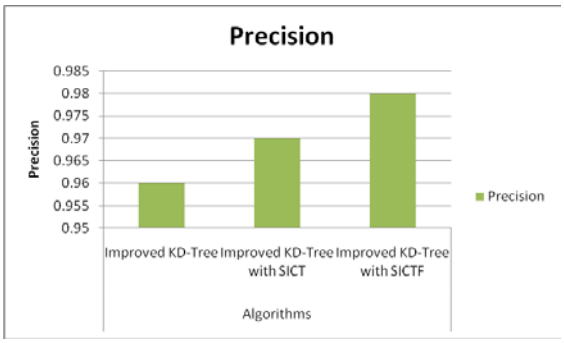

Figure 3: Precision rate comparison

The Figure 3 shows that the precision rate of Improved KD-Tree algorithm and proposed SICT with Improved KD-Tree algorithm, SICTF with Improved KD-Tree Algorithm. From this graph we can say that, our proposed algorithm SICTF with Improved KD-Tree achieves a higher level of the precision value than the other algorithm. The precision rate of Improved KD-Tree algorithm is 0.96 and the proposed system SICT achieves 0.97 level of precision and SICTF with Improved KD-Tree achieves 0.98 level of precision. From this we can say that our proposed SICTF with Improved KD-Tree 0.02 better than the existing algorithms.

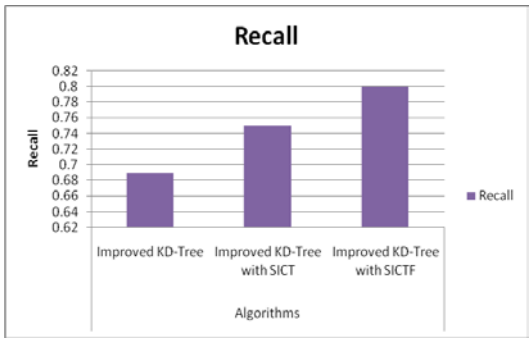

Figure 4: Recall rate comparison

The Figure 4 shows that the recall rate of Improved KD-Tree algorithm and proposed SICT with Improved KD-Tree algorithm, SICTF with Improved KD-Tree Algorithm. We measure the recall value in % at Y-axis as algorithm and consider the dataset in the X-axis. From this graph we can say that, our proposed algorithm SICTF with Improved KD-Tree achieves a higher level of the recall value than the other algorithm. The recall rate of Improved KD-Tree algorithm is 0.69 and the proposed system SICT achieves 0.75 level of recall and SICTF with Improved KD-Tree achieves 0.80 level of recall. From this we can say that our proposed SICTF with Improved KD-Tree 0.11 better than the existing algorithms.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

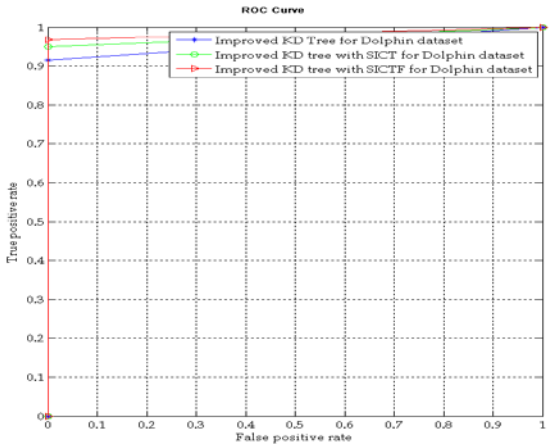

Figure 5: ROC Curve comparison

The Figure 5 shows that if the amount of Data is improved then the true positive and false positive rate also improved in the existing algorithm but when the number of number of Data is improved the true positive and false positive rate is reduced in proposed algorithms. The precision rate Comparison of the algorithm Improved KD-Tree, SICT with Improved KD-Tree and SICTF with Improved KD-Tree algorithm find most relevant sample selection. Finally our proposed algorithm achieves a higher level of the F measure value than the other algorithm.

6. CONCLUSION

Thus we have proposed two main algorithms for Sybil Identification. SICT algorithm with Improved KD-Tree used to calculate the connection threshold values; if the connection establishment is exceed the threshold then the node is identified as Sybil node. Next we proposed SICTF algorithm with Improved KD-Tree, where the maximum variance of connectivity, length and frequency of a node can be calculated for a particular time interval and the maximum variance with respect to connectivity, length, and frequency is said to be Sybil. After detecting Sybil, In order to gets the accurate and better community mining results we eliminate those nodes. We compared our algorithm performance with our previous algorithm called Improved KD-Tree. Experimental results show that the proposed SICTF with Improved KD-Tree algorithm performs well compared to the existing algorithms. And it provides more accurate results.

REFERENCES:

[1].Fortunato, S. Community detection in graphs. CoRR abs. 2009. 0906.0612.

[2].Zhuhua Cai and Christopher Jermaine, 2011. The Latent Community Model for Detecting Sybil Attacks in Social NetworksVLDB. 2011, Seattle, WA Copyright VLDB Endowment, ACM.

[3].Bimal Viswanath and Ansley Post, An Analysis of Social Network-Based Sybil Defenses, SIGCOMM'10, New Delhi, India.ACM. 2010. 978-1-4503-0201-2/10/08.

[4].J. Douceur. The Sybil Attack. In Proc. IPTPS'02, Cambridge. 2002.

[5].J. P. Bagrow, Evaluating local community methods in networks. J. Stat. Mech. 2008.

[6].N. Tran, B. Min, J. Li, and L. Subramanian. Sybil-Resilient Online Content Voting. In Proc. NSDI'09, Boston, MA. 2009.

[7].H. Yu, P. B. Gibbons, M. Kaminsky, and F. Xiao. SybilLimit: A Near-Optimal Social Network Defense against Sybil Attacks. In Proc. IEEE S&P, Oakland, CA. 2008.

[8].H. Yu, M. Kaminsky, P. B. Gibbons, and A. Flaxman, SybilGuard: Defending Against Sybil Attacks via Social Networks. In Proc. SIGCOMM'06, Pisa, Italy. 2006.

[9].Nguyen Tran, Jinyang Li, Lakshminarayanan Subramanian, and Sherman S.M. Chow, Optimal Sybil-resilient node admission control. INFOCOM, Proceedings IEEE. 2011. Pages: 3218 – 3226.

[10]. **Wei** Wei, Fengyuan Xu, Chiu C. Tan and Qun Li, SybilDefender: Defend Against Sybil Attacks inLarge Social Networks. NFOCOM, INFOCOM, Proceedings IEEE. 2012.Pages 1951 – 1959.

[11]. Viswanath, B., Post, A., Gummadi, K., and Mislove, A. An analysis of social network-based sybil defenses. In SIGCOMM. 2010.

[12]. G. Danezis and P. Mit. Sybilinfer: Detecting sybil nodes using social networks. In NDSS. 2009.

[13]. L. Xu, S. Chainan, H. Takizawa, and H. Kobayashi. Resisting Sybil attack by social network and network clustering. In SAINT. 2010.

[14]. H. Yu, M. Kaminsky, P. B. Gibbons, and A. Flaxman. SybilGuard: Defending Against Sybil Attacks via Social Networks. In Proc. SIGCOMM'06, Pisa, Italy. 2006.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[15]. H. Yu, P. B. Gibbons, M. Kaminsky, and F. Xiao. SybilLimit: A Near-Optimal Social Network Defense against Sybil Attacks. In Proc. IEEE S&P, Oakland, CA. 2008.

[16]. N. Tran, B. Min, J. Li, and L. Subramanian. Sybil-Resilient Online Content Voting. In Proc. NSDI'09, 2009. Boston, MA.

[17]. Walsh, K., and Sirer, E. G. Experience with an object reputation system for peer-to-peer filesharing. In NSDI'06: Proceedings of the 3rd conference on 3rd Symposium on Networked Systems Design & Implementation, USENIX Association, pp. 1–1.

[18]. Peterson, R. S., and Sirer, E. G, AntFarm: Efficient content distribution with managed swarms. In Proceedings of the 6th conference on Networked Systems Design & Implementation (NSDI), USENIX Association, 2006. pp. 1–1.

[19]. Piatek, M., Isdal, T., Krishnamurthy, A., and Anderson, T, One hop reputations for peer to peer file sharing workloads. In NSDI'08: Proceedings of the 5th USENIX Symposium on Networked Systems Design and Implementation, USENIX Association, 2008. pp. 1–14.

[20]. Key certification. In SSYM'98: Proceedings of the 7th conference on USENIX Security Symposium, Berkeley, CA, USA, USENIX Association, 1998. pp. 18–18.

[21]. Ziegler, C.-N., and Lausen, G, Propagation models for trust and distrust in social networks. Information Systems Frontiers 7, 4-5, 2005. Pages: 337–358.

[22]. Cheng, A., and Friedman, E, Sybilproof reputation mechanisms. In P2PECON '05: Proceedings of the 2005 ACM SIGCOMM workshop on Economics of peer-to-peer systems, ACM, 2005. pp. 128–132.

[23]. Fortunato, S, Community detection in graphs. CoRR abs/0906.0612. 2009.

[24]. Leskovec, J., Lang, K. J., and Mahoney, M. W, Empirical comparison of algorithms for network community detection. In WWW (2010), pp. 631–640.

[25]. L. Xu, S. Chainan, H. Takizawa, and H. Kobayashi. Resisting Sybil attack by social network and network clustering. In SAINT. 2010.