Trust-based Sybil Node Detection with Robust Seed Selection and Graph Pruning on SNS

Shuichiro Haruta, Kentaroh Toyoda, Iwao Sasase

Keio Univ. Sasase Lab.

7th IEEE International Workshop on Information Forensics and Security, Nov.19th 2015.

Do you use SNS (Social Networking Services) ?

Most of us use SNS such as Facebook or Twitter.

However, recently malicious accounts appear

Motivation for our research

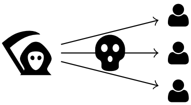
Detect malicious account called ”Sybil”, effectively

Sybils on SNS

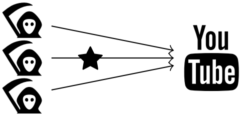
Many malicious accounts created by attackers.

Aims of Sybils:

Ex.1) Sending spam

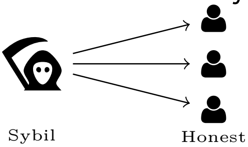


Ex.2) Illegal voting



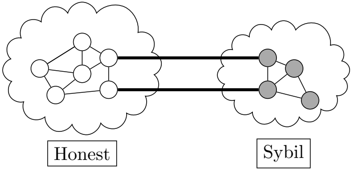
Procedures of Sybils:

- Create many accounts
- Send friend requests to ordinary accounts to attack



System Model

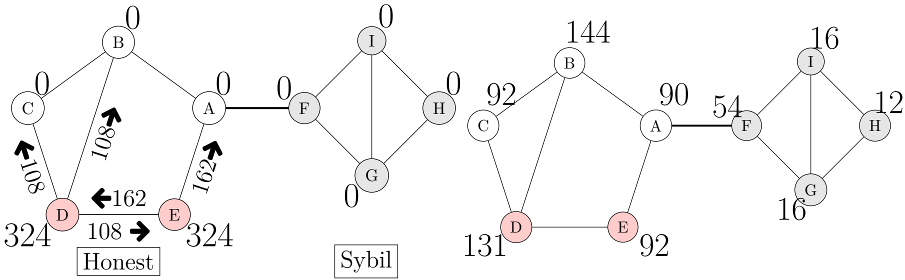
- Non-directional graph SNS (e.g., Facebook)
- We call ordinary nodes as ”honest nodes”
- Sybils can make friends among Sybils
  - Sybils pretend to be honest nodes
- Sybils try to make friends with honest nodes
  - Acceptance ratio of friend request is relatively small e.g., friend request from stranger



The relationships among Sybils and honests are called ”**AE (Attack Edge)**”.

Conventional Scheme 1 : SybilRank[1]

- Sybil detection scheme Cao et al. proposed
- repeatedly distribute trust values of non-Sybil nodes (**honest seeds**)to their neighbors
- honest nodes tend to get higher trust values
  - since trust values Sybils get are only via AE

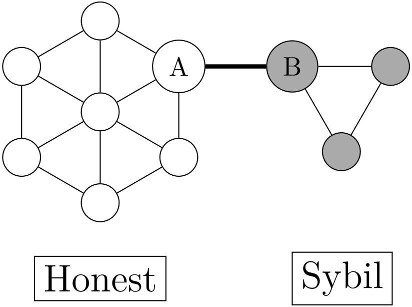


Initial state

After four iterations

Conventional Scheme 2 : Graph pruning scheme [2]

- Zhang et al. proposed **to enhance SybilRank's detection accuracy**
- Prior to SybilRank, doubtful edges are pruned
- Prune the relationships between two nodes with few common friends



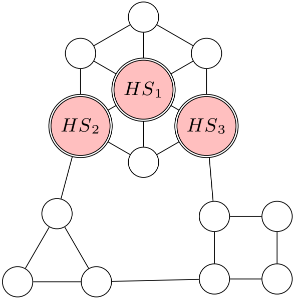
The number of common friends between A and B is 0. This thick edge is pruned.

Shortcomings of SybilRank

- SybilRank randomly chooses honest seeds from high degree nodes

↓

Honest seeds may be concentrated in specific community



- Honest nodes far from honest seeds (HS) get insufficient trust value
- it causes **false positive**

Distribution of high degree nodes

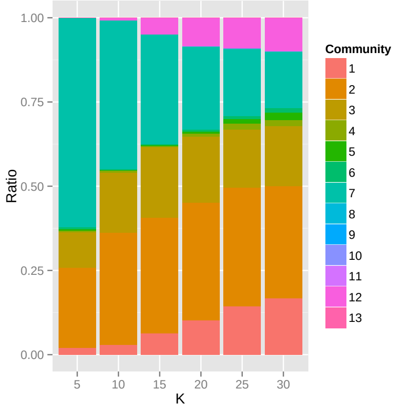


Figure: The ratio of nodes with high degree in each community.

Figure shows communities where the top K% highest degree nodes belong. (Dataset is Facebook)

Example: K $= 5$

Most of high degree nodes belong to 4 out of 13 communities.

Conventional Seed Selecting ···

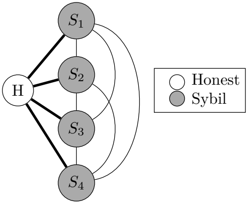
Honest seed tends to concentrate in specific community

Shortcomings of Graph Pruning Scheme

- Conventional graph pruning scheme[2] prunes AE based on the number of common friends

↓

Attacker can easily increase the number of common friends and in that case, pruning AE is difficult



If an attacker creates relationships, the number of common friends between honest and Sybil is increased.

The common friends of H **and** S_1 **are** $\mathrm{S}_2, \mathrm{S}_3, \mathrm{S}_4$**.**

Proposed Scheme: Seeds Selecting Scheme

- Select honest seeds from each community
- We use community detection scheme

We select seeds uniformly from entire graph

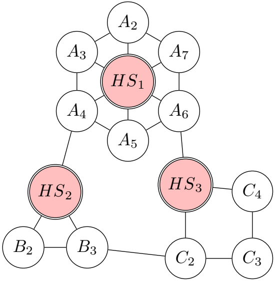


Figure: **Prop. Seed Selection**

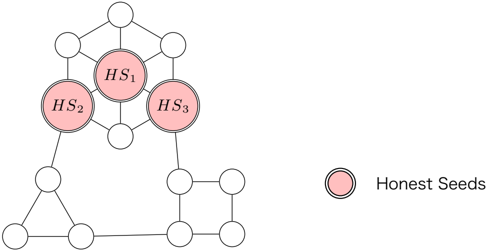
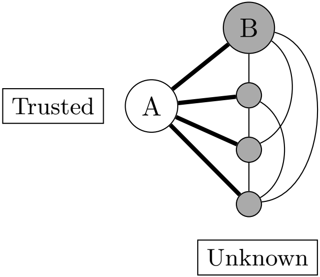


Figure: Conv. Seed Selection

Proposed Scheme: Graph Pruning Based on Density of Relationships

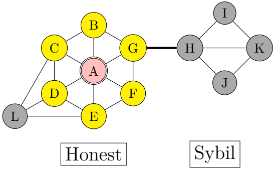
- Consider how many trusted relationships exist
- Prune relationships between a trusted node and an untrusted node with probability



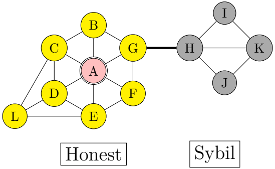
- Node A has already been regarded as trusted
- Node B is unknown and has 4 friends
- $\frac{1}{4}$ of node B's friends are trusted
- Pruning probability between A and B is $P_{\text{prune}}(A, B) = 1 - \frac{1}{4} = \frac{3}{4}$

Calculating Trusted Area

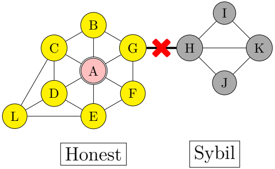
We define nodes with dense relationships as TA (Trusted Area). We calculate TA with three steps.



Trust neighbors of seed



Include nodes to TA based on density of relationships.



Prune doubtful edges

Simulation Settings : Attack Scenario

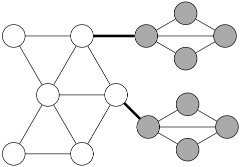
Attack scenario 1

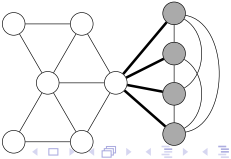
- Conventional attack scenario
- Add totally 200 AEs from randomly chosen Sybils to Honest nodes

Attack scenario 2

- Add totally 200 AEs in below style
- Honest nodes that accept friend requests from strangers are limited

In the simulation, we intensively increase the common number of friends





Simulation Setting : 2

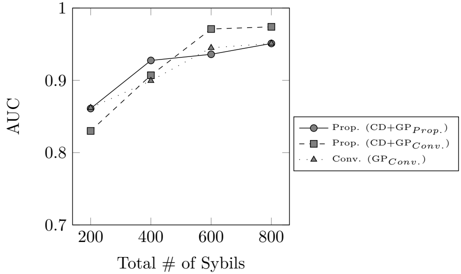
- Evaluating the AUC (Area Under Curve) of ROC (Receive Operating Characteristic) curve[3].
- AUC indicates ability of a classification algorithm and ranges 0 to 1.

Table: Parameter values used in the simulation.

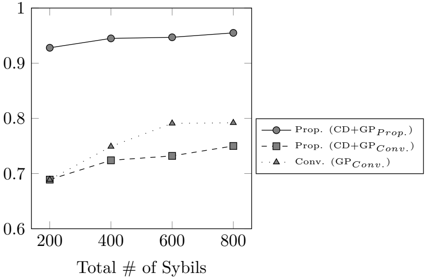
parameter	value
dataset	Facebook [4]
number of nodes	4039
number of edges	88234
graph model	BA model[5]
n_{att}	5
K	10
simulation tools	R with igraph[6]

The Sybil Detection Accuracy: AUC vs total number of Sybils

Attack scenario 1



Attack scenario 2

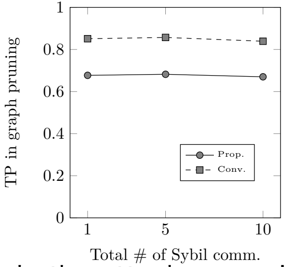


- In the attack scenario 1, the both proposed schemes achieve almost same accuracy with the conventional scheme
- In the attack scenario 2, the proposed scheme with community detection and our graph pruning considerably improve AUC against the conventional scheme

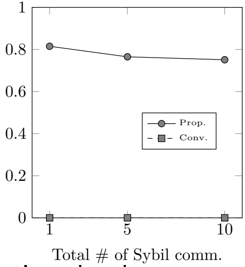
True Positive Rate in Graph Pruning vs Total number of Sybil Comm.

We define **TP** (True Positive) rate as **how accurately AEs are pruned**

Attack scenario 1



Attack scenario 2

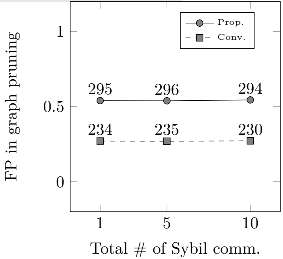


- In the attack scenario 1, the conventional scheme effectively prunes AE. The proposed scheme degrades.
- In the attack scenario 2, while the conventional graph pruning scheme cannot prune AE, the proposed scheme prunes AE with high TP.

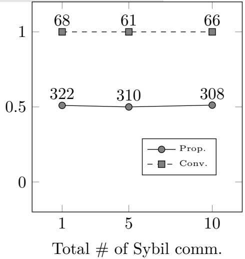
False Positive in Rate in Graph Pruning vs Total of Sybil comm.

We define **FP** (False Positive) rate as **the ratio that the non-AE are inaccurately pruned**[1].

Attack scenario 1



Attack scenario 2



- In the attack scenario 2, the conventional scheme cannot accurately prune AE.
- Although FP seems to be high in both scheme, it can be acceptable since the # of entire edges is about 90,000.

[1] The numeric above a point is the average number of pruned

Conclusion

- We have proposed a Sybil nodes detection scheme with robust seed selection and graph pruning on SNS
- The first proposal is seed selecting scheme by detecting communities and choosing seeds from them
- The second one is a graph pruning scheme that considers trusted area
- By computer simulation, we show that our scheme achieves high detection accuracy even if attackers make a large number of common friends

Citations

[1] Q. Cao, M. Sirivianos, X. Yang, and T. Pregueiro, "Aiding the detection of fake accounts in large scale social online services," USENIX NSDI, pp.2–15, USENIX Association, 2012.
[2] H. Zhang, C. Xu, and J. Zhang, "Exploiting trust and distrust information to combat sybil attack in online social networks," in Trust Management VIII, pp.77–92, 2014.
[3] J.A. Hanley and B.J. McNeil, "The meaning and use of the area under a receiver operating characteristic ROC curve.," Radiology vol.143, no.1, pp.29-36, 1982.
[4] J. Leskovec and J.J. Mcauley, "Learning to discover social circles in ego networks," Advances in neural information processing systems, pp.539–547, 2012.
[5] A.L. Barabasi and R. Albert, "Emergence of scaling in random networks," science, vol.286, no.5439, pp.509-512,1999.
[6] G. Csardi and T. Nepusz, "The igraph software package for complex network research," InterJournal, p.1695, 2006.

Thank you for listening.