基于女巫攻击预防技术的电信网络诈骗团伙识别

任泽华

2020年12月14日

报告内容

1 研究背景

2 研究内容

3 结论展望

研究背景

■ 社交网络安全现状

随着通信技术的革新与互联网的兴起，网络诈骗愈发猖獗。网络诈骗总是呈现出团伙诈骗的形式，这与人们的信任模式有很大的关系：

人们或多或少拥有从众的心理。

■ 现有的社交网络诈骗检测方法

- 利用有向图模型的欺诈账户检测（图分析）。
- 针对用户行为特征的欺诈检测（机器学习等方法）。
- 大规模社交网络社区的检测方法（团伙社区发现）。
- P2P网络的欺诈团伙检测方法（分布式网络检测）。

女巫攻击：单节点对外表现出多重身份的攻击行为。此模式和诈骗团伙极为相似，若迁移到当前情景下进行诈骗团伙的检测，将对于解决这一问题做出极大贡献。

■ 女巫攻击

微软研究院的Douceur教授在2002年的文章《The Sybil Attack》中首先提出，被认为是基于对等（P2P）网络的一种基本攻击形式。

它的主要表现方式为：**单一节点具有多重身份标识**，通过控制系统中大部分节点来达到削弱网络冗余性、降低网络健壮性、破坏网络正常活动、盗取其他节点个人信息等目的。

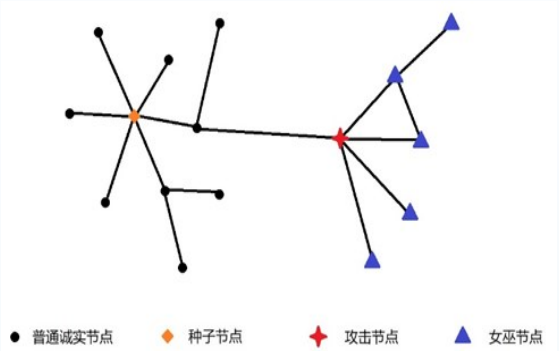


图1 社交网络图谱的基本模型

■ 模型概述

每个节点表示一个账户，每条边表示两个账户建立了信任关系。

节点分为普通**诚实节点**、**种子节点**（被确定为诚实的节点）、**攻击节点**（具有真实身份的恶意节点）、**女巫节点**（虚假身份的节点）。

研究背景

■ 攻击方式

➢ 孤立攻击

一名攻击者将自己的计算与存储资源分配给自己控制的女巫节点，以通过连接到其他节点来获得他人承认。

① 一名攻击者连接到一些种子节点，并创建一些sybil节点；
② 一名攻击者连接到一些排名最高的诚实用户，并创建一些sybil节点；
③ 一名攻击者连接到一些诚实用户，并创建一些sybil节点；
④ 一名攻击者连接到排名最高的诚实者之一，并创建多个sybil群组；
⑤ 种子节点尝试创建一些sybil节点；
⑥ 诚实节点尝试创建一些sybil节点。

➢ 协同攻击

多名攻击者将自己的计算与存储资源分配给其控制的女巫节点，攻击者能够以任何方式彼此联系。

① 一组或多组攻击者试图连接到某些种子节点并创建一些sybil节点；
② 一组或多组攻击者试图连接到一些顶级诚实用户，并创建一些sybil节点；
③ 一组或多组攻击者试图连接到某些诚实用户并创建一些sybil节点；
④ 一群攻击者试图与一些最诚实的诚实分子建立联系，并创建多个sybil群组；
⑤ 一组或多组种子节点尝试创建一些sybil节点；
⑥ 一组或多组诚实节点尝试创建一些sybil节点；
⑦ ……

报告内容

1 研究背景

2 研究内容

3 结论展望

研究内容

■ **研究内容部分主要包括：**

➢ **基于随机游走的女巫节点检测算法：**

本部分主要介绍了本文使用的三种随机游走算法： SybilGuard、 SybilLimit和SybilRank的基本原理。

➢ **针对不同类型网络的实验验证：**

本部分主要比较了中心化、多中心、分布式网络下三种算法的检测时间、检测精度等指标。

➢ **基于节点行为的检测（原算法改进）：**

本部分利用节点行为信息，使用KNN和朴素贝叶斯结合的方法对检测算法进行了改进。

基于随机游走的女巫节点检测算法

■ SybilGuard

➢ **该算法基于以下几点基本假设：**

（1）社交网络正在快速混合（节点正在快速于其他节点建立联系）。

（2）初始验证的节点一定是诚实节点（种子节点）。

（3）恶意用户可能会创建多个节点，但Sybil社区相对独立于诚实社区。

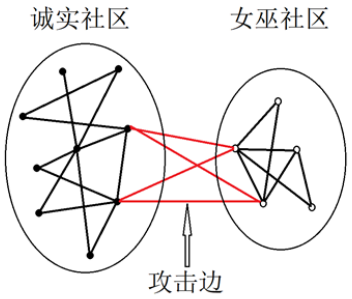


图2 女巫攻击网络图谱示意

➢ **算法基本步骤：**

（1）选择一个初始节点V，保证这个节点为诚实节点（Seed），以此为基础进行随机游走；

（2）设该节点的度为d，从该节点引出d条随机路线，每遇到一个节点时都执行随机选择，每条路线的长度为w；

（3）待测节点S也通过随机游走的方式进行路线选择，如果两个节点走出的路线有超过某一阈值t的数量相交时，就认为S是诚实节点，否则就认为它是Sybil节点。一般取t=d/2。

基于随机游走的女巫节点检测算法

■ SybilLimit

➢ **与SybilGuard的主要区别：**

（1）交叉点条件：每个节点使用r个长度为w的随机路径进行并行游走，而不是一次一条。从而降低了攻击者利用搜索间隙发动攻击的可能性。 $r = o(\sqrt{n})$ 、 $w = o(\log n)$

（2）平衡条件：SybilLimit依靠其新的平衡条件来解决原方法中的未命中路线。

（3）参数估计技术：使用一种新的参数估计技术，将Sybil节点和诚实节点放在一起估计，保证了无论攻击者行为如何，都不会过高估计参数r：同时游走条数。

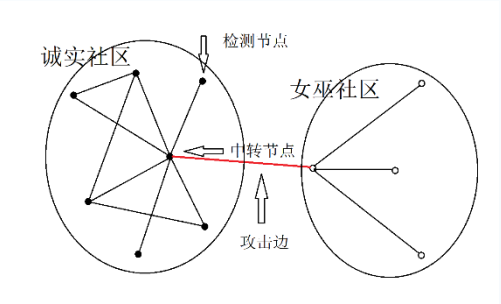


图3 诚实社区与Sybil社区相对独立

表1 两种方法的性能比较

攻击边数g（协议未知）	接受节点数量	
	SybilGuard	SybilLimit
$o(\sqrt{n}\log n)$	$o(\sqrt{n}\log n)$	$o(\log n)$
$o(\sqrt{n}\log n)$ 到$o(n\log n)$	无限制	$o(\log n)$
小于15,000	2000左右	10左右
大于15,000 且小于100,000	无限制	10左右

基于随机游走的女巫节点检测算法

■ SybilRank

➢ 算法基本步骤：

（1）通过w=O(logn)幂次迭代，信任概率从已知的非Sybil节点（信任种子）流出，在整个网络中扩展并偏向非Sybil区域；

（2）根据节点的程度归一化信任度对其进行排名；

（3）在排名列表中选取前几名的用户，并向全体用户公告它们被怀疑为Sybils账户。

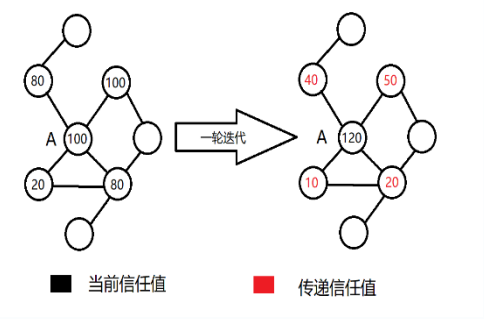


图4 信任传递模型

➢ 算法主要思想：

在随机游走的基础上加入排序。首先将随机游走进行幂次迭代，把正常节点的信任值分配给其他节点，并进行标准化，最后根据标准化结果排序。

$O(\mathrm{n}\log n)$——比前两种略高

➢ 算法优势：

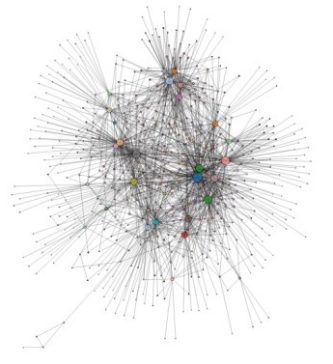
计算成本不会随其使用的信任种子节点数的增加而增加。这有助于使用多个种子来提高系统对种子选择错误的鲁棒性。

针对不同类型网络的实验验证

■ 使用的网络简介

我们使用Network Repository 网站上面的开源网络数据集进行测试，使用Facebook上面的三个真实网络图谱数据Caltech36、Reed98和American75进行了测试。

➢ 网络拓扑图：

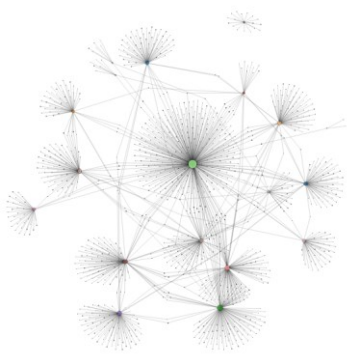


Caltech36

769 nodes,
16.7k edges

Reed98

962 nodes,
18.8k edges



American75

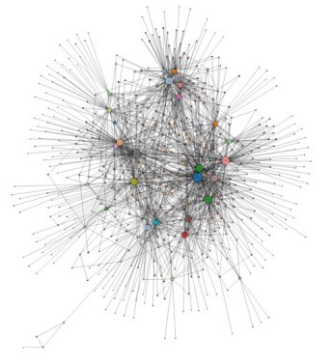
6.4k nodes,
217.7k edges

针对不同类型网络的实验验证

■ 使用的网络简介

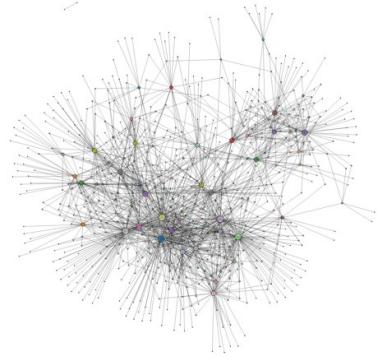
我们使用Network Repository 网站上面的开源网络数据集进行测试，使用Facebook上面的三个真实网络图谱数据Caltech36、Reed98和American75进行了测试。

➢ 网络拓扑图：



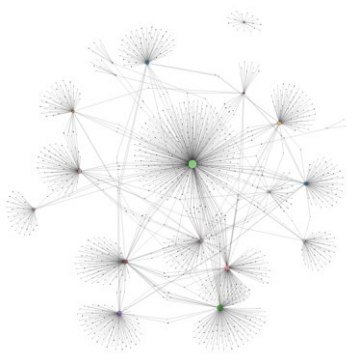
Caltech36

中心化



Reed98

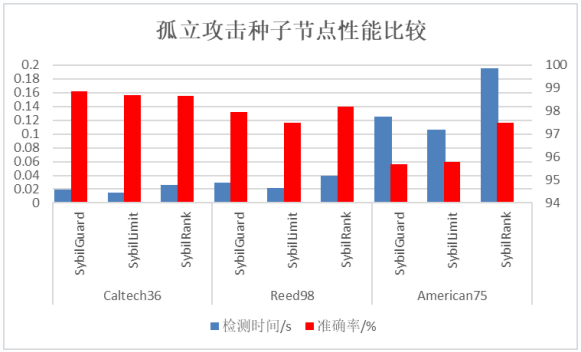
多中心

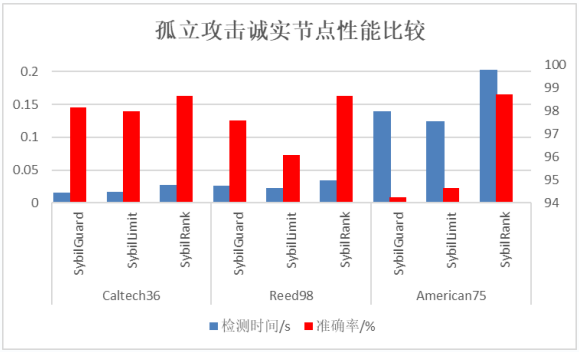


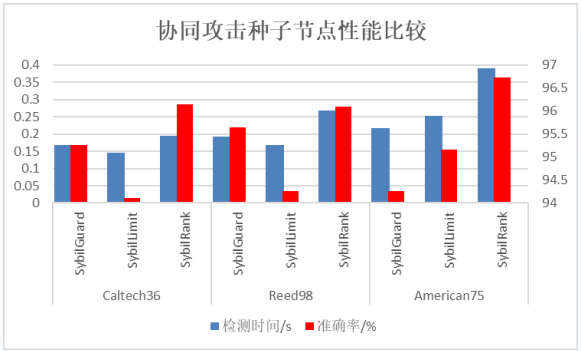
American75

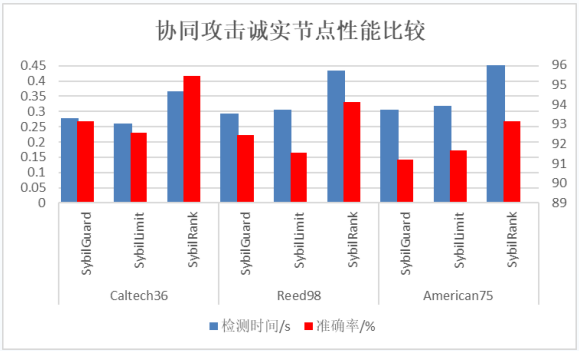
分布式

针对不同类型网络的实验验证









基于节点行为的检测

■ 检测模型

➢ 相关系数计算：

$$R_{X,Y} = \frac{\sum(X-\bar{X})(Y-\bar{Y})}{\sum(X-\bar{X})^2\sum(Y-\bar{Y})^2}$$

我们使用斯坦福大学的SNAP数据库中2012年推特用户行为的真实数据Higgs Twitter Dataset，对每一对行为属性进行相关性分析，得到了如表所示的分析结果。从表中我们可以看出最具相关性的一对属性为粉丝数和关注数。

表7 相关系数比较表

属性	粉丝数	关注数	点赞数	状态
粉丝数	1			
关注数	0.8982	1		
点赞数	-0.2081	0.1296	1	
状态	0.5826	0.4963	0.0281	1

基于节点行为的检测

- **平均F/F率：**

$$F(s) = \frac{1}{n}\sum_{i=1}^{n} F(X_i)$$

F(x)为样本节点x在属性F下的值（如F1(x)表示x的粉丝数）

s为数据集

- **算法流程：**

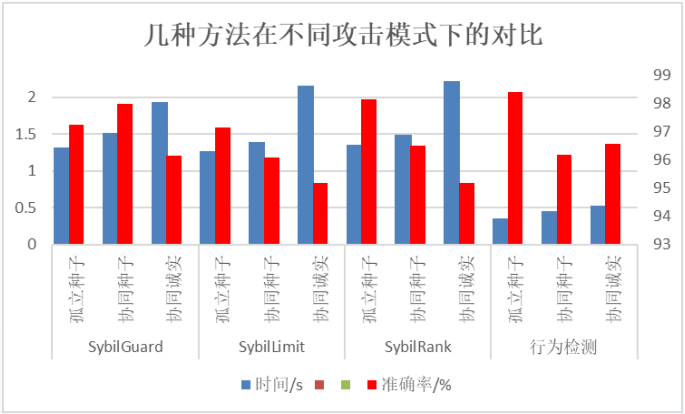
① 计算平均F/F率；

② 如果0.5F(S)＜F(X)＜1.5F(S)，使用KNN算法；

③ 否则使用朴素贝叶斯算法；

④ 返回分类结果。

- **实验结果：**



报告内容

1 研究背景

2 研究内容

3 结论展望

结论展望

- **结论：**

对于中心化的网络，几种算法精度差别不大，而此时SybilRank要花费更多的时间；而对于分布式网络，SybilRank在牺牲一些检测时间的基础上拥有较好的检测结果。所以我们在选用算法时需要根据网络结构灵活决定。

在此基础上，我们引入用户行为，对用户进行了画像，使用KNN和朴素贝叶斯等机器学习分类手段对用户进行了判别。与之前的网络图检测算法相结合，算法效率与准确度取得了较大的提高。

- **展望：**

接下来的工作中，我们打算进一步深化用户行为检测的算法。在进一步调研中我们发现Zhou Q等人使用受害者预测模型来提高检测精度，这是一个全新的角度，站在被攻击对象的角度建模，估计哪些用户是潜在的被攻击对象，及时向他们发出危险警报，这在未来的女巫攻击团伙检测中将是一个全新的发展方向。

谢谢!

参考文献

[1] Douceur J R. The Sybil Attack[C]. international workshop on peer to peer systems, 2002: 251-260.

[2] Davis C R , Fernandez J M , Neville S , et al. Sybil attacks as a mitigation strategy against the Storm botnet[C]// International Conference on Malicious & Unwanted Software. IEEE Computer Society, 2008.

[3] Mohsen K. BrightID Anti-Sybil[EB/OL]. [2020.11.06].https://github.com/BrightID/BrightID-AntiSybil

[4] Valarmathi M L, Meenakowshalya A, Bharathi A, et al. Robust Sybil attack detection mechanism for Social Networks - a survey[C]. international conference on advanced computing, 2016: 1-5.

[5] Gunturu R. Survey of Sybil Attacks in Social Networks.[J]. arXiv: Cryptography and Security, 2015.

[6] Garg R, Sharma H. Prevention Techniques for Sybil Attack[C]. international conference on bioinformatics, 2012, 11(10): 3060-3064.

[7] Yu H, Kaminsky M, Gibbons P B, et al. SybilGuard: defending against sybil attacks via social networks[C]. acm special interest group on data communication, 2006, 36(4): 267-278.

[8] Yu H, Gibbons P B, Kaminsky M, et al. SybilLimit: A Near-Optimal Social Network Defense against Sybil Attacks[C]. ieee symposium on security and privacy, 2008: 3-17.

[9] Cao Q , Sirivianos M , Yang X , et al. Aiding the Detection of Fake Accounts in Large Scale Social Online Services[C]// Usenix Conference on Networked Systems Design & Implementation. USENIX Association, 2012.

[10]谈磊, 连一峰, 陈恺. 基于复合分类模型的社交网络恶意用户识别方法[J]. 计算机应用与软件, 2012(12):1-5.

[11] Zhou Q, Chen G. An Efficient Victim Prediction for Sybil Detection in Online Social Network. [J]. IEEE Access. 2020: PP. 1-1. 10.1109/ACCESS.2020.3007458.