针对 Sybil 攻击的防御技术的发展

任泽华
（西安交通大学，电信学部，西安，710049）

摘要：本文针对分布式网络中最常遇到的网络攻击：Sybil 攻击的相关防御技术的发展进行了深入介绍。叙述了 Sybil 攻击的名称来源和各种已有的、潜在的威胁。对 Sybil 攻击进行了简要的分类，即按照通信方式、身份来源、同时性三个方面来分。对比了不同种类的 Sybil 攻击对于分布式网络各种协议的危害程度。在具体介绍防御算法时主要分了三个阶段：基于资源测试和安全证书的防范方法、基于社交网络图谱的 Sybil 节点检测方法和基于信任模型和节点行为的 Sybil 攻击检测方法。这三个阶段是随着攻击技术的进步、网络规模的扩大和网络功能的加强应运而生的。针对第一阶段，我们对于基于资源测试的节点检测方法进行了重点阐述；针对第二阶段，我们对 SybilGuard、SybilLimit 和 SybilDefender 三种方法进行了重点介绍；针对第三阶段，我们对基于模糊逻辑推理的信任评估方法进行了重点复述。最后，我们对于 Sybil 攻击的有利应用进行了简要分析，并且分析了针对 Sybil 攻击的攻防博弈未来的发展方向，得出了其将朝向与人工智能与大数据结合的方向发展的结论。

关键词： Sybil 攻击 分布式网络 社交网络 网络拓扑 模糊逻辑推理

Development of defense technology against Sybil attacks

Ren Zehua

(Xi'an Jiaotong University, School of Telecommunications, Xi'an 710049）

Abstract: This paper gives an in-depth introduction to the development of the Sybil attacks which is most commonly encountered in distributed networks. Describes the name source of the Sybil attack and the potential threats. The simple classification of Sybil attacks is based on three aspects: communication method, identity source, and simultaneity. Compares the degree of damage of different types of Sybil attacks against various protocols in P2P networks. In the specific introduction of the defense algorithm, there are mainly three stages: the prevention method based on resource testing and security certificates, the Sybil detection method based on the social network graph and the Sybil attack detection method based on the trust model and conventional behavior. For the first stage, we focused on the node detection method based on resource testing; for the second stage, we focused on the methods of SybilGuard, SybilLimit and SybilDefender; for the third stage, we focus on the trust evaluation method based on fuzzy logic reasoning. Finally, we briefly analyzed the beneficial applications of the Sybil attack. The future direction of the attack and defense game against Sybil attacks is also analyzed. It is concluded that it will develop in the direction of combining with artificial intelligence and big data.

Keywords: Sybil attack; Distributed network; Social network; Network topology; Fuzzy logic inference

0 引言

微软研究院的 Douceur 教授在 2002 年的一篇论文《The Sybil Attack》【1】中首先提出了 Sybil 攻击的概念。Sybil 攻击，又被译作女巫攻击，它的名称来源是同名小说改编的电影《Sybil》，主要讲述了一个具有 16 重人格的女人的故事，而 Sybil 攻击正是以她的名字来命名的。它被认为是基于对等网络（P2P 网络）进行攻击的一种基本形式。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。就像它名称来源的电影情节，这种“一人分饰多角”的攻击会给网络带来前所未有的破坏与难以估量的损失。

现实生活中有许多网络是基于 P2P 结构而构建的，有些网络可能拥有一个或多个中心节点，而有些网络中每一个节点的身份都是平等的。这种网络常见的形式有：以比特币和以太坊为代表的区块链网络、在物联网环境下各种传感器组成的无线传感器网络、基于人际关系建立的人与人之间的社交网络等等。它们可能在网络功能、通讯方式、安全协议等等方面有着较大的区别，但它们都有一个共同特点，那就是每个普通节点相对独立、处理信息上采取分布式的方法、往往采取冗余的方式来抵抗可能存在的网络攻击。对于以区块链为代表的 P2P 网络，它们往往没有一个可信任的中间节点，攻击者在利用一个主机获取大量用户身份后，可以对整个网络发动包括 51%攻击、日蚀攻击、DDos 攻击等等形式的攻击，来达到伪造交易记录、窃取数字货币的目的。针对以无线传感器网络为基础的物联网应用中，Sybil 攻击会对整个网络的路由分配、资源调度和不良行为检测等方面造成破坏【2】，在物联网设备日益普及的今天，这种攻击可能会对我们的生命财产带来严重威胁。针对社交网络应用中，Sybil 可能会通过伪造大量虚假用户，甚至有可能会影响某些大型网络投票的结果【3】，在诸如互联网商务中基于用户评价的信用评分体系可能会受到严重篡改从而造成信任危机【4】。

本文第一节将介绍 Sybil 攻击的主要分类，其中包括三个维度：通信方式、身份来源、同时性，并比较了几种分布式网络的基础协议受到 Sybil 攻击的程度。第二节首先从最基础的 Sybil 攻击防范手段开始介绍，也就是 Sybil 攻击提出者 Douceur 教授提出的节点资源测试方法与带中央节点控制的颁发安全证书的方法，在这个基础上产生了随机密钥预分配、位置验证法和分布式安全证书颁发的方法。第三节主要介绍了基于社交网络的 Sybil 节点检测方法，它的主要思想是将 P2P 网络抽象为社交关系图，而 Sybil 节点往往会与诚实节点呈现不同的拓扑结构，通过网络拓扑的检测即可确定 Sybil 节点。其中比较重要的方法有 SybilGuard、SybilLimmt 和 SybilDefender。第四节将从恶意节点检测的最新方法上入手来介绍当前研究的主要方向，即基于节点行为和信任模型的恶意节点检测法。其中介绍的主流方法的思路是采用非线性灰度预测模型来预测直接信任关系，并且通过模糊逻辑推理的方法将信任关系、节点行为等等评价标准综合起来，模拟人的判断方式来达到检测效果。第五节会对上述方案进行总结与评估，并分析未来可能的发展方向。

1 Sybil攻击的分类与各种协议受攻击的程度

1.1 主要分类

我们将 Sybil 攻击主要按以下三个维度来进行分类【2】：

1.通信方式

直接通信：Sybil 节点与诚实节点进行直接沟通，在二者相互通信时，恶意设备会进行侦测，同样，Sybil 节点发送的消息实际上是经由恶意设备发出的。

间接通信：诚实节点不能与 Sybil 节点直接沟通，只能通过恶意设备进行诱导，使得诚实节点发送到 Sybil 节点的信息路由到恶意节点上，而在物理上，这些 Sybil 节点是不存在的。

2.身份获取方式

虚假身份：当身份创建不需要进行认证或者认证方式容易破解时，恶意设备可以通过直接构建虚假身份的方式来创建 Sybil 节点。

盗取身份：当攻击者无法伪造身份时（比如地址空间受到限制），他们会通过攻击或策反某些合法节点从而把它们变成 Sybil 节点，完成身份的获取。

3.节点进攻是否同时

同时：攻击者让他所创建的 Sybil 节点同时接入网络，虽然一个设备可能同一时刻只能模拟一个身份，但是通过轮询的方式可以同时对外呈现多个身份。

非同时：攻击者在某一段时间内显示大量的身份，而在某个确定的时间内只显示少数身份，这种

方式使得 Sybil 节点定位变得困难。

1.2 各种协议受攻击的程度

分布式存储协议：Sybil 攻击会破坏对等存储系统中的复制和碎片化机制，影响数据的存储。

路由协议：对于多路径分散路由来说，看似不相交的两条线路可能会经过同一设备创建的不同 Sybil 节点，从而泄露信息；对于地理路由来说，Sybil 节点可能会一次出现在多个位置，影响路由选择。

信息汇总协议：如果攻击者拥有足够多的 Sybil 节点，他就能左右系统的信息收集机制。

投票协议：同上，足够多的 Sybil 节点可以左右投票结果。

资源分配协议：攻击者凭借大量的 Sybil 节点占用了大量不平等分配的系统资源。

不良行为检测协议：Sybil 节点会串通一气，误导诚实节点对于他们的不良行为的检测。

表一给出了不同形式的 Sybil 攻击对各种协议的危害程度，可以看出大部分 Sybil 攻击方式都会对我们提到的大部分协议产生危害，所以 Sybil 节点的检测与去除就显得至关重要。

	通信方式		身份获取方式		同时性	
	直接	间接	虚假	盗取	同时	非同时
分布式存储协议	√	√	√	√	√	
路由协议	√		√	√	√	
信息汇总协议	√	√	√	√	√	
投票协议	√	√	√	√	√	√
资源分配协议	√	√	√	√	√	√
不良行为检测协议	√	√	√	√	√	√

表一：不同形式的 Sybil 攻击对各种协议的危害程度

2 基于资源测试和安全证书的Sybil攻击防范方法

2.1 基于资源测试的 Sybil 节点检测法

早在 Douceur 教授提出 Sybil 攻击概念时，他就提出了两种预防 Sybil 攻击的方法【1】，它们分别是本小节介绍的基于资源测试的 Sybil 节点检测法和下一小节介绍的基于中央节点颁发安全证书的防范法。第三小节主要分析基于这两种方法的后续研究进展。图一为该分布式网络的基本模型，每个节点叫做一个实体，而恶意实体可以对外呈现出多个不同的实体身份，本地实体在与其他实体共享信息时必须判断它是善意的还是恶意的。

基于资源测试的检测方法原理非常简单，其中分为直接验证和间接验证。因为攻击者一般是不可能和整个网络的资源相抗衡的，换句话说就是它拥有的计算、通信、存储资源都是有限的。所以验证节点可以通过让被验证节点执行单台主机无法完成的计算任务，或者让它存储一些单台主机无法存放的大型数据，或者利用通信资源的限制，让分散的各个节点同时接收数据，若其中有 Sybil 节点，它们必然要通过同一条通信线路接收消息，就会产生冲突。而节点的间接验证方式是节点相信它绝对信任的节点所担保的其他节点，而这也就在一定程度上节约了因为验证而消耗的大量资源。

这种方法看似完美，但在内部机制上却存在着致命的漏洞。Douceur 教授在自己的论文中就证明了四个定理，证明了直接验证方式和间接验证方式都不可避免地会受到 Sybil 攻击。

直接验证：

定理 1：即使在资源严重受限的情况下，恶意节点也可以伪造恒定数量的多个身份。

定理 2：每个正确的节点必须同时验证其所提供的所有身份；否则，恶意节点可以伪造无限数量的身份。

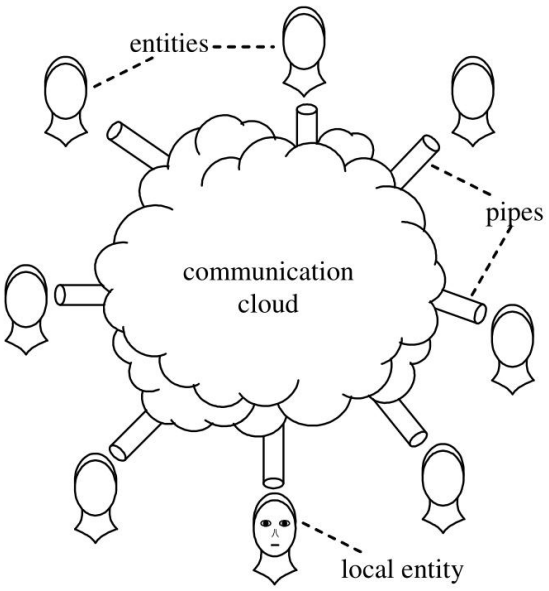
间接验证：

定理 3：足够多的恶意节点集可以伪造无限数量的身份。

定理 4：系统中的所有节点必须同时执行其身份验证；否则，恶意节点可以伪造恒定数量的多个身份。

造成这些漏洞的主要原因之一是直接验证的不同步性，使得攻击者可以在不同的时间内分别冒充不同的节点，从而理论上对外可以呈现无限的身份；

而另一个重要原因是若其计算资源大于最小计算资源的多倍时，就可以毫无阻碍地同时扮演这么多个身份。而间接验证就在于 Sybil 节点之间是可以互相担保的，这可以追溯到古罗马时期的拜占庭将军问题。同一时刻网络中一定会有 Sybil 节点不被诚实节点发现而存在，而当 Sybil 节点达到了一定的程度，它们会欺骗网络中所有的节点，网络中可以容错的最大 Sybil 节点数大致为总数的 1/3。



图一：分布式网络的基本模型

2.2基于中央节点颁发安全证书的防范法

这种思想其实正在被普遍应用于日常生活中，比如说银行交易系统、学校考勤系统、飞机订票系统、政务办理系统等等。它们的共同特征是有一个大家共同信赖的中央机构，也就是中央节点，这个节点可以对每一个节点进行担保，为每个通过验证的用户节点颁发安全证书。有人认为这样就失去了分布式网络的根本属性，其实我们这里的提到的中央节点权限可能不大，仅仅承担节点诚实性的担保任务，而用户节点之间相互交流、协同工作时，中央节点是不参与的。这样看似能从根本上避免 Sybil 攻击，然而事实真的是这样吗？

首先，这种方式对于中央节点的依赖过大，每次建立连接前的验证都要通过中央节点，如果中央节点被攻击，整个网络将会陷于瘫痪。而即使拥有了一个类似于中央信托机构的节点，攻击者仍然可以通过破解中央节点的安全认证协议来达到攻击的目的，到了这时，中央机构就不是安全的防卫者，而变成了攻击者的帮凶，它会为攻击者提供官方的、“可靠的”安全证书，而此时整个网络就会面临崩溃。

2.3 一些新的研究进展

在以上两种思想的基础上，研究者们发展出了一些新的方法，其中有代表性的是随机密钥预分配法、位置验证法【2】【5】和分布式的安全证书颁发法【6】。

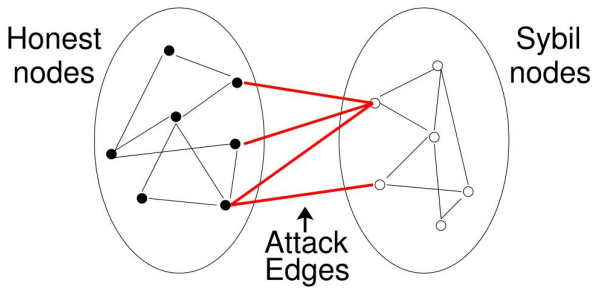
随机密钥预分配法：将节点身份与分配的随机密钥相关联，密钥用于节点之间的私密通信和节点的身份验证。其中主要有三个发展阶段。（1）基本密钥池的方法，也就是随机密钥之间分配，这会造成无法进行节点到节点的独立身份验证；（2）单空间成对密钥分配方法（将唯一的密钥分配给每对节点），这种方法的问题是网络大小受到密码空间的限制，一旦攻击者成功捕获了超过一定数量的节点，整个空间就会受到威胁，并且他可以制造任意数量的身份；（3）多空间成对密钥分配方法（设置服务器随机生成一个 m 个密钥空间的池，每个池都有唯一的私有信息。每个节点在 m 个键空间内选择 k 个，如果两个相邻节点共有一个或多个密钥空间，则它们可以使用相应的单个空间方案来计算其成对秘密密钥。这种方法看似完美，却会大量占用内存和运行资源。

位置验证法：这被认为是一种比较有前途的方法，假设网络一旦部署就无法移动。在这种方法中，网络将验证每个节点的物理位置。可以使用此方法检测到 Sybil 节点。移动攻击者可能能够通过在一个位置被验证为一个身份，然后移动到另一个位置并被验证为另一个身份，来呈现多个身份。为了克服这种攻击，可以同时验证所有节点的位置。另外，给定攻击者移动性的上限，只需要同时测试特定范围内的节点即可。但是这种方法仅限于网络节点位置不变的网络，比如特定情境下部署的传感器网络，适用范围有限。

分布式的安全证书颁发法：这种方法是集中式网络和分布式网络的折衷，将一批“元老节点”作为证书的颁发机构，只有获得了大部分“元老节点”认可的节点才能被整个网络认可。这里也存在着一些问题，比如说让本该一个节点完成的任务重复多次，造成了一定的资源消耗，但是制作方法使得安全证书颁发机构具有一定的容错性且不易被攻击。

3 基于社交网络图谱的Sybil攻击检测方法

虽然上述方法在检测 Sybil 节点方面取得了巨大的贡献，但是随着社交网络的快速兴起，人们的日常生活越来越依靠于它；而物联网技术的发展也在要求着更加快速、健壮的无线传感器网络；数字货币的逐渐兴起，特别是今年央行试发行数字货币，加上区块链技术在各行各业的广泛应用，人们对于分布式网络的性能提出了更高的要求。而传统的检测方式无不需要大量的计算资源和复杂的密码算法，这无法满足我们对于网络快速性、容错性和减小 Sybil 节点的综合要求。此时，基于社交网络图谱的 Sybil 节点检测方法就应运而生。如图二所示，它将网络抽象为社交网络图谱，节点代表网络节点，边代表二者建立的信任关系（或叫好友关系）。其中将节点分为两部分：诚实节点和 Sybil 节点，同种且相邻近的节点构成了“节点社区”。由于图中只有这两种节点，我们易证，它们之间必然会形成一个二分图，其中连接两部分节点的边叫做攻击边（Attack Edges）。



图二：社交网络图谱的基本模型

3.1 SybilGuard 和 SybilLimit 法

1. SybilGuard

Haifeng Yu【7】等人在 2006 年提出了一种基于社交网络检测 Sybil 节点的算法，SybilGuard 法，这也是基于社交网络图谱的第一种检测方法。这是一种完全分散的检测算法，所有操作均针对给定节点，它保证了诚实节点以极大的可能接受其他的诚实节点，并且被其他诚实节点接受；还保证了诚实节点只能接受有限数量的 Sybil 节点。该算法基于以下几点基本假设：（1）社交网络正在快速混合，即诚实节点正在快速地于其他诚实节点建立联系（2）初始验证的节点一定是诚实节点，基于这个初始诚实节点来检验其他节点。（3）恶意用户可能会创建多个节点，但是它们能够说服诚实节点接受它们的概率比较小，换句话说就是攻击边相对较少，这就导致了 Sybil 社区相对独立于诚实社区，这就为我们基于网络拓扑检测 Sybil 节点提供了可能。

SybilGuard 法的基本步骤为：

（1）选择一个初始节点 V，保证这个节点为诚实节点，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。一般去 t=d/2。

SybilGuard 法得以实现主要基于以下几点理论：

（1）Sybil 社区相对独立于诚实社区，从 Sybil 节点出发的路径到达诚实社区的可能性较小，有时仅为一两条边（如图三所示）。

（2）回路只能在路线的起点形成。

（3）随机路线沿相同方向多次穿越某个边缘一次（即循环），或者进入 sybil 区域，则它是有问题的。

（4）若两条路径相交，那么接下来它们将相重合。

关于游走长度的选择也是有讲究的：w 的值必须足够小，以确保（i）验证者的随机路径完全保持在诚实区域内的可能性很高；（ii）sybil 社区的大小不宜过大。另一方面，w 必须足够大，以确保路由以高概率相交。而此处我们选取 $w = o(\sqrt{n}\ \log n)$比较合适。对于任何连接且非双向的社交网络，从统一随机诚实节点开始的长度为 w 的随机游走将遍历任何 g 个攻击边缘的概率的上限为 gw/n。特别是当 $w = o(\sqrt{n}\ \log n)$时，该概率为 $o(1)$。

2. SybilLimit

同样是 Haifeng Yu【8】等人在 2008 年对自己的算法提出了改进，形成了 SybilLimit 算法。如表二所示，它可将 Sybil 攻击值降低为 $o(\ \log n)$，在大型网络中，这种优势尤其明显。它的主要思想为：在初始化阶段，每个节点都构建自己的路由表。在联机阶段，可疑节点会将观众的标识符和地址发送给验证者节点，该节点将比较可疑者列表中的证人以尝

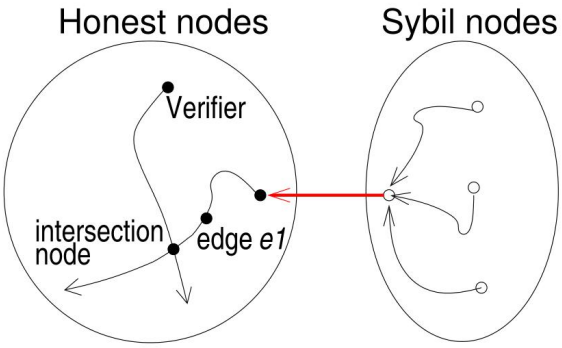
试查找匹配项。如果验证者在两个集合中找到匹配项，它将要求两个集合中具有相同身份的观众验证可疑节点的身份，并根据此过程的结果确定是否接受或拒绝该节点。如果两组之间没有交集，则验证程序将中止并拒绝该节点，将其标记为攻击者。

SybilLimit 区别于 SybilGuard 的主要方面为：

（1）交叉点条件：在 SybilLimit 中，每个节点使用$r = o(\sqrt{n})$个长度为$w = o(\log n)$的随机路径进行游走，而不是像之前一样一次只进行一条线路的游走。从而降低了攻击者利用搜索间隙发动攻击的可能性。在任何给定的情况下，对手都可以伪造 w 条跨越攻击边缘并进入诚实区域的长度为 w 的不同随机路线。此处 SybilLimit 通过使用比 SybilGuard 小得多的 w 来减少此类路由的数量。

（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：它使用一种新颖的参数估计技术，将 Sybil 节点和诚实节点放在一起估计，这就保证了无论攻击者行为如何，都不会过高地估计参数 r，即同时游走的路线条数。



图三：诚实社区于 Sybil 社区相对独立

攻击边缘数量 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

表二：两种方法的性能比较

3.2 SybilDefender 法

Wei wei 等人在 2013 年提出了一种全新的方法 SybilDefender，这种方法的主要包括一个用于识别 sybil 节点的 sybil 识别算法，一个用于检测 sybil 节点周围的 sybil 社区的 sybil 社区检测算法，以及通过降低一定精度以大大提高检测速度的近似算法。这就使得此算法具有了检测 Sybil 社区的能力，因此能够快速定位大量的 Sybil 节点。

SybilDefender 法主要由以下两个部分组成：

1. Sybil 识别算法

用一个诚实节点检验另一个可疑节点，用随机游动的方法。以均等的概率进行游走。这一点类似于前面的两种方法 。

（1）输入社交图谱 G、可靠节点 h，输出识别阈值：首先执行 f 个短暂的随机游走，长度为 logn，得到 f 个均匀分布的节点，它们高概率是诚实节点，与源节点组合形成验证节点。针对每个验证节点，进行随机游走，长度在[lmin，lmax]区间范围内遍历一遍，验证 R 个节点。每个长度输出 f+1 个值，计算均值、标准差并输出。

（2）先进行一次随机游走，长度 l 大于（1）中的最小值，把值不小于 t 的拿出来和（1）中的均值比较。这些值减去（1）中均值，如果大于标准差×20，输出该节点是 Sybil 节点，否则将长度 l 加倍迭代，直到 l 大于 lmax，此时输出该节点不是 Sybil 节点。

2. Sybil 社区检测算法

用检测到的 sybil 节点确定 sybil 区域。其基本理论依据是：sybil 节点的部分随机游走往往被困在 sybil 区域内。

（1）估计（2）中使用的部分随机游走所需的长度。给定一个初始长度 l，计算走不到 l 的路径的比例，若小于 0.95，l 加倍，直到这个比例大于 0.95，输出 l 的值。

（2）定义了一种度量，d 是集合 S 中所有节点的度之和，a 是在 S 中具有一个端点而在一个端点上的边的数量。用 a/d 表示相连程度的大小。可以用贪婪算法来计算。首先执行从已知 sybil 节点 s 发出的 R 个部分随机游走，其长度由算法 3 决定。然后按其频率按降序对所有遍历的节点进行排序。对排序后的

列表进行迭代，如果度量不增加，将遇到的节点添加进来，检查排序后的列表中的所有节点后，算法记录当前连接程度值，从列表顶部开始新的迭代，并检查不在 S 中的每个节点。重复此过程，直到程度值在两次连续迭代的末尾保持相同为止。算法输出 S 作为检测到的 sybil 社区。

文中还提到了一种牺牲部分检测精度以提高检测速度的算法，Combo 算法：如果我们找到一个 sybil 节点，我们将直接分析识别方法中得出的简单随机游走。该分析基于连接程度度量。我们按照节点的频率，简单随机游动遍历的次数以降序对节点进行排序，然后将节点迭代添加到检测到的 sybil 集中，直到连接程度值在两次连续迭代的末尾保持相同。这种方法虽然会稍微降低精度，但却会大大节约运算时间。

3.3 其他同时期的方法

1. Sybil Infer：【10】是一种中心化检测方法，,利用贝叶斯不确定推理理论,计算节点为 Sybil 的置信度.该方法能够取得较低的漏警率,但计算复杂度高。总体时间复杂度为 $o(\sqrt{v}\ \log v)$，其中 v 是网络中的节点数。 Sybil Infer 仅适用于包含 30K 节点的网络，不适用于较大的网络。

2. Symon：【11】在 Symon 中，每个节点都有一个称为 Symon 的非 Sybil 节点。在网络中动态选择了非 Sybil 节点或 Symon。每个 Symon 都有监视网络行为的责任。

3. Sybil Shield：【12】是一种比较新颖的方法，用于划分具有多个小型，大型和中型社区的社交网络图。Sybil Shield 比起 Sybil Defender 提高了诚实节点的接受率。不仅使用随机路由方法来检测 Sybil，还使用代理漫游方法。

3.4 几种检测方法比较

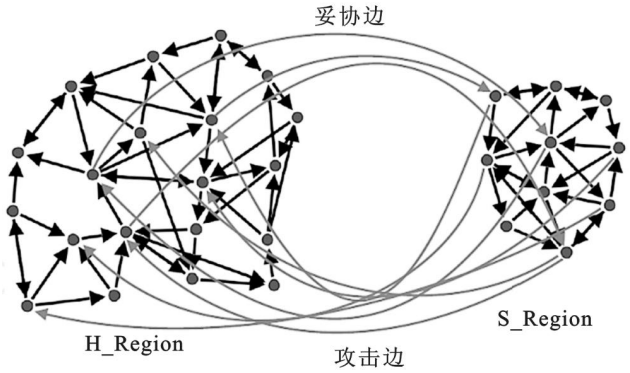
表三将几种基于社交网络图谱的检测方法之间的优劣展现了出来，其中 Sybil Defender 法是最成功，也是应用最广泛的方法，而 Sybil Guard 和 Sybil Limit 方法是最基础的，最容易解释此类算法思想的基础算法。

3.5 最新发展

近几年，国内研究者对于基于社交网络图谱的 Sybil 节点检测方法也有所研究，其中比较有代表性的是王永程等【13】提出的针对有向社交网络的 Sybil 检测方法 SybilGrid。对于有向图而言，他们创造性地定义了于攻击边对应的妥协边，作为反向关系，如图四所示。由于是有向图，所以环路数量大大减小，而“妥协边“的定义也使得误判为 Sybil 节点的概率大大减小。但是这种方法也要局限性，那就是仅限用于单向关注的社交网络，比如说微博和推特，而对于双向联系的社交网络，比如说微信和 Facebook，这种方法就可能无法工作。

序号	算法名称	使用计算方法	优点	缺点
1	Sybil Guard	随机游走	提高检测精度, 接近理想状态	一次只检测一个节点
2	Sybil Limit	随机游走	可靠性提高	仅在快速混合的网络中有效
3	Sybil Defender	限制节点的游走	效率高、可量化	初始节点为 Sybil 会失败
4	Sybil Infer	基于贝叶斯定理的随机游走	比 Sybil Guard 更可靠	计算过多
5	SyMon	动态选择	性能快速混合	大型网络中变得不适用
6	Sybil Shield	使用源节点测试方案	很好地降低误判率	节点增加，误判率上升

表三：各种基于社交网络检测方案比较



图四：针对有向社交网络的 Sybil 攻击模型

4 基于信任模型和节点行为的Sybil攻击检测方法

在上述模型中，都仅仅将网络抽象为社交图谱，而使用了网络图谱的拓扑关系来进行判断，随着攻击技术的发展，这种方法也面临着挑战。由于现在的许多分布式网络结构非常复杂，计算复杂度也随之提高，而且在网络拓扑结构不知道的情况下这些方法便无法进行，所以研究者们提出了一类新的方法:基于信任模型和节点行为的 Sybil 攻击检测方法。

4.1 基于模糊逻辑的信任评估方法

Hamdi Ouechtati【14】等人在 2020 年提出了一种基于模糊逻辑推理的信任评估方法，这也是这一领域最新的检测方法，接收到有关请求访问的对象的推荐时，推荐管理器首先根据发送者和推荐对象之间的现有社会关系对推荐进行分组。然后，它基于以下因素评估推荐值之间的相似性：（1）内部相似度，表示每个对象发送的每个推荐值与其所属社区的推荐值相比的相似度；（2）外部相似度，代表每个对象发送的每个推荐值与其他社区的推荐值之间的相似度。

计算相似性和社交关系程度的目标是检测碰撞攻击的出现并评估收到的推荐的信任度。主要分为以下几个部分：

1. 推荐分类

OOR：由与推荐对象具有所有权关系的几个对象发送的推荐集；C-LOR：由与推荐对象具有共置关系的几个对象发送的推荐集；C-WOR：由与推荐对象具有 C-WOR 关系的几个对象发送的推荐集；SOR:由与推荐对象有社会关系的几个对象发送的推荐集

2. 评估推荐之间的相似性

推荐管理器将每个对象发送的每个推荐值进行比较：（i）与从其所属社区接收到的推荐的中位数，以评估其内部相似度，以及（ii）与中值的平均值其他社区，以评估其外部相似性。（具有相同社会关系的对象构成社区）。这使得我们可以在总体水平上评估内部和外部相似性之间的相似性。例如，如果内部相似度很高，而外部相似度很低，那么推荐值可信度就很低。

3. 社会关系程度的评估

推荐管理器将推荐值可信度与社交关系的程度相结合，以便在推荐的信任级别评估中考虑社交关系的程度。因此，社会关系的程度越强，信任度越高。

4. 推荐书的信任度评估和共谋攻击的检测

我们使用模糊推理评估收到的建议的信任级别，并检测碰撞攻击的出现。使用名为非常低，低，中，高和非常高的隶属函数表示内部相似度，外部相似度和推荐值可信度。

先将内部、外部相似度合成为推荐值可信度（RVC），再将推荐值可信度（RVC），社会关系程度（DSR）合成为信任级别。基于模糊逻辑的信任评估包括两个主要阶段：（i）评估收到的建议的信任级别；（ii）检测共谋或单一攻击以过滤不适当的建议。这种方法也取得了很好的效果，能够达到 94％左右的精度，在最坏的现实情况下，当 40％的对象发送错误的建议时，能够保持几乎相同的精度，大约为 91％。

4.2基于非线性多元灰色预测模型预测直接信任程度的方法

Hui Xia【15】等人 2019 年用基于非线性多元灰色预测模型来预测特定对象的直接信任，考虑到信任概念的模糊性和不确定性，同时引入了模糊逻辑方法来综合这些信任元素。基于以下假设：人们倾向于信任更熟悉的人，更喜欢在态度，兴趣，背景和个性方面与自己相似的朋友。此方法将信任分为两类：熟悉信任（FT）和相似性信任（ST）。通过直接信任（DT）和推荐信任（RT）来计算熟悉度信任；基于外部相似性信任（EST）和内部相似性信任（IST）计算相似性信任。其中最关键的是提出了一种基于内核的非线性多元灰色预测模型，用于预测特定对象的 DT，称为 KGM（1，n），它克服了以

前的灰色模型的局限性。最终通过模糊逻辑的方法综合各种因素，达到了优良的效果。

5 总结与展望

本文主要回顾了 Sybil 攻击的防御技术的发展历程，以从刚提出 Sybil 攻击时产生的资源测试、安全证书方法为第一阶段，其后又衍生出了分布式的安全证书法等方法；而随着攻击技术的发展于社交网络的兴起，基于社交网络图谱的 Sybil 节点检测方法产生，其主要思想是根据社交网络图谱的拓扑结构来定位 Sybil 节点；而随着网络的不断扩大和网络拓扑变得不可知，又产生了基于信任模型与节点行为的综合检测法，此方法综合利用了其他学科的技术，比如说模糊逻辑推理和灰色预测模型，将本来难以量化的行为与信任程度等等因素综合起来，形成了比较好的检测结果。

而 Sybil 攻击也并不是完全有害，利用得当它也可以为我们所用。任何技术都是具有两面性的，sybil 攻击也不例外。虽然它在 P2P 网络、社交网络和物联网应用的网络中带来了不可估量的损失，但是它也可以被应用于对抗其他网络攻击。Carlton R. Davis【16】等人就提出了一种利用 Sybil 攻击来对了 storm 僵尸网络的方法。当僵尸网络中的大部分节点都替换为 Sybil 节点，处于我们的掌控中时，我们也就消灭了这个僵尸网络。所以说，sybil 攻击的攻防博弈的研究是很有意义并且很有前景的，它不仅仅帮助我们抵抗可能存在的 sybil 攻击，也让我们在更加了解 sybil 攻击的基础上让它为我们所用，达到检测与对抗其他网络攻击的目的。

在未来，我认为 Sybil 攻击的攻防博弈有着广阔的发展空间，一方面是黑客创造假身份的技术不断提高，另一方面是我们的网络规模不断扩大、网络结构不断复杂、网络功能不断增加，对于新应用场景下的 Sybil 攻击还会继续进行。而对于分布式网络来说，Sybil 攻击是进行其他攻击，比如说 Eclipse 攻击、DDos 攻击等等的基础，攻击者正是能够以多重虚假身份骗过我们，才能进行下一步的更有危害的攻击。而随着 5G 技术的发展，Sybil 攻击必然会在这一新的土壤上继续生根发芽，在未来我们还有很长的路要走。而未来的发展方向在我看来，未来的一段时间内是会继续向着基于行为和信任模型的检测来进行。而随着技术、攻击方式和需求的进一步变化，Sybil 攻击的检测方式可能会朝着基于人工智能、大数据等等平台的更有效的检测方式发展。而那时，攻击方式也必然向着同一方向前进，这也正是彼此矛盾的双方对立又统一，相互促进的关系。

参考文献

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

[2] Newsome J, Shi E, Song D, et al. The Sybil attack in sensor networks: analysis & defenses[C]. information processing in sensor networks, 2004: 259-268.

[3] Tran N, Min B, Li J, et al. Sybil-resilient online content voting[C]. networked systems design and implementation, 2009: 15-28.

[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] Wei W, Xu F, Tan C C, et al. SybilDefender: A Defense Mechanism for Sybil Attacks in Large Social Networks[J]. IEEE Transactions on Parallel and Distributed Systems, 2013, 24(12): 2492-2502.

[10] Danezis G, Mittal P. SybilInfer: Detecting Sybil Nodes using Social Networks.[C]. network and distributed system security symposium, 2009.

[11] Jyothi B S, Janakiram D. SyMon: A practical approach to defend large structured P2P systems against Sybil Attack[J]. Peer-to-peer Networking and Applications, 2011, 4(3): 289-308.

[12] Shi L, Yu S, Lou W, et al. SybilShield: An agent-aided social network-based Sybil defense among multiple communities[C]. international conference on computer communications, 2013: 1034-1042.

[13] 王永程, 孟艳红. 针对有向社交网络的 Sybil 检测方法[J]. 西安电子科技大学学报, 2016, 43(02):199-204.

[14] Ouechtati H, Azzouna N B, Said L B, et al. A Fuzzy Logic Based Trust-ABAC Model for the Internet of Things[C]. advanced information networking and applications, 2019: 1157-1168.

[15] Xia H, Xiao F, Zhang S, et al. Trustworthiness Inference Framework in the Social Internet of Things: A Context-Aware Approach[C]. international conference on computer communications, 2019: 838-846.

[16] 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.