第35卷 第15期 计 算 机 工 程 2009年8月
Vol.35 No.15 Computer Engineering August 2009
·安全技术· 文章编号：1000—3428(2009)15—0121—03 文献标识码：A 中图分类号：TP393.08

P2P网络中Sybil攻击的防御机制

胡玲玲，杨寿保，王 菁
(中国科学技术大学计算机科学与技术学院，合肥 230026)

摘 要：针对P2P网络中的Sybil攻击，基于小世界模型提出一种防御机制。将Sybil攻击团体的发现归结为最大流/最小割问题，引入虚拟节点并利用爬行器找到Sybil攻击团体。实验结果证明，该机制能将Sybil攻击者和P2P系统分开，减少系统中Sybil攻击节点所占比例。
关键词：P2P技术；Sybil攻击；最大流；最小割

Defense Mechanism for Sybil Attack in P2P Network

HU Ling-ling, YANG Shou-bao, WANG Jing
(School of Computer Science and Technology, University of Science and Technology of China, Hefei 230026)

【Abstract】Aiming at the Sybil attack in P2P network, this paper proposes a defense mechanism based on the small world model. The discovery of Sybil group boils down to max-flow/min-cut problem, introduces virtual node and uses a crawler to find Sybil group. Experimental results show that this mechanism can isolate the Sybil attacker nodes from P2P system, and reduce Sybil nodes' percentage in the system.
【Key words】P2P technology; Sybil attack; max-flow; min-cut

1 概述

P2P系统的开放性和匿名性导致其容易受到攻击。随着P2P系统规模的扩大，将出现越来越多的恶意节点。学者已针对Sybil攻击展开了很多研究[1-2]。一个信任的中心认证机构通过为每个人分配唯一的标识符，可以有效阻止Sybil攻击。但该机制需要中心认证机构同时为每个标识符付出一定代价，且在很多情况下不适用，例如很难选择或建立整个系统都信任的独立实体。中心认证机构会成为一个瓶颈，并可能引起QoS攻击。

文献[3]采用EigenTrust算法减少P2P网络中非授权文件的共享，从网络中隔离恶意节点，但没有解决Sybil攻击问题。本文研究并分析社会网络，提出一种新的Sybil攻击防御机制。

2 Sybil攻击

在具有可信任认证中心节点的P2P网络环境中，一个未知远程节点不可能呈现不同身份。但在没有逻辑中心和可信任认证中心节点的完全分布式P2P网络里，节点和节点间进行一对一的通信时，无法对对方身份进行认证，导致未知远程节点可以表现为具有不同身份。如果本地节点和未知远程节点没有直接的物理连接，本地节点就会把远程节点仅看作信息的抽象(身份)。P2P系统必须保证不同身份属于不同节点，否则，当一个本地节点错误选择了一个恶意主机进行文件下载时，恶意主机会冒充好节点向它发送非法数据，导致好节点的信誉值降低，网络中的资源将被浪费。P2P系统里仿造多个身份的攻击被称为Sybil攻击。该攻击产生的主要原因是用户无需任何代价就可以创建新的身份和节点。攻击者利用此漏洞在网络中传输非法文件，从而破坏系统中文件的安全性，消耗网络中节点的连接资源而无需担心自己会受影响。

3 Sybil攻击防御系统的设计

在本文采用的社会关系网络图中，节点代表身份；边代表人为的信任关系。连接诚实团体和Sybil攻击团体的边是Sybil攻击边。本文设计的基本想法如下：如果恶意节点创建了太多的Sybil攻击身份，则关系网络图的节点分布会受到很大影响，即出现最小割，将Sybil攻击节点和整个系统分离。直接寻找上述最小割很困难，因为无法获取整个网络的拓扑结构和每条边的端点。即使可以获得整个网络的拓扑，寻找最小割仍然是一个NP问题。

3.1 最大流/最小割问题

对于一个给定的有向图 $G=(V, E)$，其中每条边$(u, v)\in E$均有一个非负容量 $c(u, v)\geqslant 0$。G中存在2个特殊点，源点s和汇点t，利用s和t之间不同边的容量找到从s到t的最大流即为s-t最大流问题。网络G中分离s和t的一个边的集合称为G的一个割集，割集容量最小者称为网络G的最小割。

Ford和Fulkerson提出的最大流/最小割理论证明了寻找一个网络的最大流问题等同于划分s和t的最小割问题。因此，本文基于最大流/最小割理论给出基于最小割问题的社会团体的定义。

定义1 如果$[S, T]$是图$G(V, E)$的最小割，且$s\in S, t\in T$，则S是相对于t的一个社会团体。

最大流问题的多数解决方案都基于如下假设：整个图的结果可以很容易地获得。该假设不适用于动态P2P系统，因为P2P系统的整体结构随着节点的自由加入和离开不断变化。本文选择最短增广路径算法找出P2P系统中源点和汇点间的最短路径，因为它只要利用系统中部分拓扑结构就可以有效解决最短路径问题。

基金项目：国家自然科学基金资助项目(60673172)；国家“863”计划基金资助项目(2006AA01A110)
作者简介：胡玲玲(1983－)，女，硕士研究生，主研方向：对等网络；杨寿保，教授、博士生导师；王 菁，博士研究生
收稿日期：2009-02-21 **E-mail：**lnghu@mail.ustc.edu.cn

3.2 Sybil攻击团体

3.2.1 理想状态

本文定义无向图中的通信，其中每条边都具有单位容量。在理想状态下，每条边都被用于计算且不具有平凡割。

定义 2 Sybil 攻击团体是一个节点集合 $G' \subset E$，对于集合中的所有节点 $v \in G'$，v 在 G'中具有的连接数至少等于它和 $(V\text{-}G')$中节点所具有的连接数。

定义 3 给出变量 s'，代表 s 和集合$(G'\text{-}s)$中节点间的连接数。变量 t'代表 t 和集合$(V\text{-}G'\text{-}t)$中节点间的连接数。

结论 寻找 P2P 系统中的 Sybil 攻击团体 G'，可以转化为寻找 G 中以 s 为源点，t 为汇点的 s-t 最小割问题，其中，s' 和 t'要大于最小割的大小。当最小割的边被删除后，所有对于 s 可达的点都在 Sybil 攻击团体里。

图 1 给出了将 Sybil 攻击团体和整个系统分割开的例子。在结论中给 s'和 t'的条件是为了避免平凡割问题。但一个 Sybil 攻击团体会包括很多类似的平凡割。为了解决上述问题，本文引入一个虚拟节点，并选择多个种子节点和虚拟节点相连。其中，种子节点和虚拟节点之间的容量为无穷大。

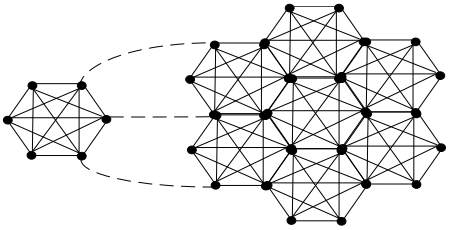

图 1 基于最小割的 Sybil 攻击团体

3.2.2 现实环境中 Sybil 攻击团体的发现

在构造 P2P 系统时，可以基于 Jon Kleinberg 模型，因为很多相关研究[4-5]表明，P2P 网络拓扑是符合小世界模型的，所以节点的入度和出度可以遵循 power-law 分布。

在理想状态下寻找 Sybil 攻击团体的主要问题是需要快速获取很多节点的入度和出度信息。图 2 描述了本文爬行器如何寻找属于 Sybil 攻击团体的节点。先选择一些种子节点，如图 2 所示。然后爬行器从种子节点开始获得它们的入度和出度。用此方法可以获得图 2 中的集合 c。当集合 c 中的节点被确定后，可以获得它们的出度。这些出度信息可以被分为 2 类，即从集合 c 到集合 c 以及从集合 c 到种子节点。所有其他出度都被看作从集合 c 到集合 d，而集合 d 中的节点最终汇聚为一个虚拟节点。

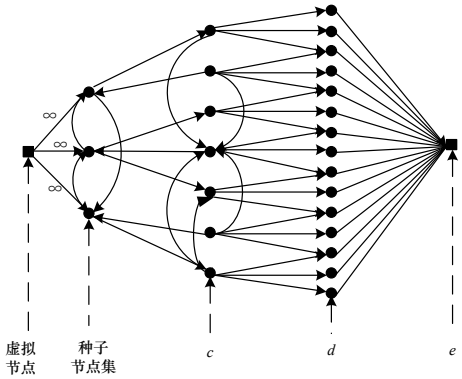

图 2 Sybil 防御系统中爬行器的工作原理

为了使用最小割算法将 Sybil 攻击团体和 P2P 系统分割开，本文采用一个虚拟节点连接所有节点，如图 3 所示。假设为每个节点都建立到虚拟节点之间的无向边，边的容量为 α，则图 G 变为扩展图 G_α。扩展图 G_α 中的 α 对于寻找 Sybil 攻击团体起着重要作用。其值与系统聚类的数目有密切关系。

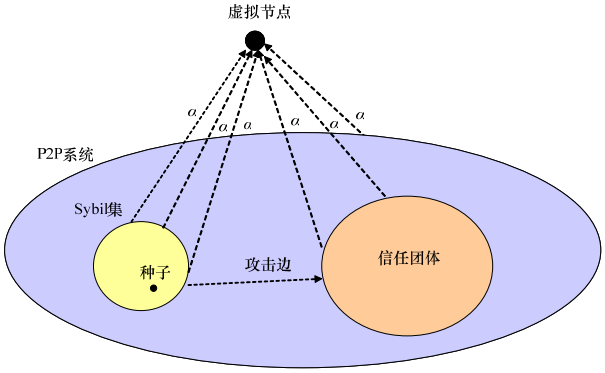

图 3 引入虚拟节点的 P2P 系统

当 α 为 0 时，G 中所有节点与 t 之间的最小割都为割(t, V)，将 t 和所有 G 中的节点分离。此时，最小割算法将整个图 G 聚为一类。

当 α 的值趋于无穷大时，最小割算法将 G 中所有的 n 个节点划分为 n 个类(即 n 个团体)，每个团体只包含其自身。

当 α 的值在上述 2 种极限情况之间变化时，G 的 n 会被聚成 1 个~n 个团体，聚类的形式主要取决于 G 的结构和边的容量分布。团体的数目随着 α 的增加是非递减的。

P2P 系统中寻找 Sybil 攻击团体的具体算法如下(其中，$G=(N, E)$；节点数为 p；$q \in N$)：

```
// N 为系统中节点的数目，E 为边的数目
While number of iterations is less than desired do
  Set k equal to the number of peers in seed set,
    Perform maximum flow analysis of G,
      Get the group G'.
    Identify non-seed peers, p*∈G';
      With the highest in-degree relative to G;
  For all r∈G' such that in-degree of r equals p*
    Add r to seed set;
    Add edge (p, r) to E with infinite capacity.
  End for
  Identify non-seed peers, s*;
 With the highest out-degree relative to G;
For all s∈G' such that out-degree of s equals s*
  Add s to seed set s;
    Add edge (p, s) to E with infinite capacity.
    End for
    Re-crawl so that G uses all seeds.
  Let G reflect new information from the crawl.
End while
End
```

先利用最大流算法产生信任团体。通过分析历史信息获取网络的连接状态，利用最短增广路径算法产生信任团体 G。然后选择种子，增强团体并通过爬行器不断扩充种子集合。

当 Sybil 攻击团体被确定后，爬行器会把消息传播到系统中的其他节点。当一个节点选择另一个节点进行交易时，它会先判断对方是否属于 Sybil 攻击团体，再决定是否从对方下载文件或将文件共享给对方。本文通过将 Sybil 攻击团体和 P2P 系统分开，保护整个系统的安全性。模拟实验证明了基于小世界模型的 Sybil 防御机制的有效性。

4 模拟实验

假设节点到达系统遵循泊松分布，节点的生存周期呈指

数分布，即在本文描述的系统中，大部分节点只生存很短时间。这和实际 P2P 应用情况相符，因为在现实 P2P 的应用中，多数用户下载完所需文件后都会离开，即“搭便车”[6]。在该情况下，只有很少一部分节点会长时间停留在系统中继续提供服务。

本文实验基于 P2P 文件共享系统，并假设 3 种场景：没有攻击者，一个攻击者和 4 个攻击者。假设攻击者节点有无穷大的生命周期，因为它们一般不会离开系统，且不考虑恶意节点出现故障的情况。Sybil 攻击者的主要目标是尽可能多地获取系统中的 IDs，因此，假设获取一个新的 IDs 意味着新节点的加入。实验使用 C++编写源码，相关参数见表 1。

表 1　基于小世界模型的 Sybil 攻击防御系统参数

参数	定义
N	系统中节点的数目
N_{sybil}	Sybil 攻击者的数目
k	种子的个数
α	节点和虚拟节点之间的权值
R_a	合法节点的到达率
L	一个合法节点的平均生命周期

给出没有防御机制情况下，P2P 系统中受 Sybil 攻击的情况。在系统达到稳定状态后，Sybil 节点开始攻击。无防御系统时，Sybil 攻击的节点数与时间的关系如图 4 所示。

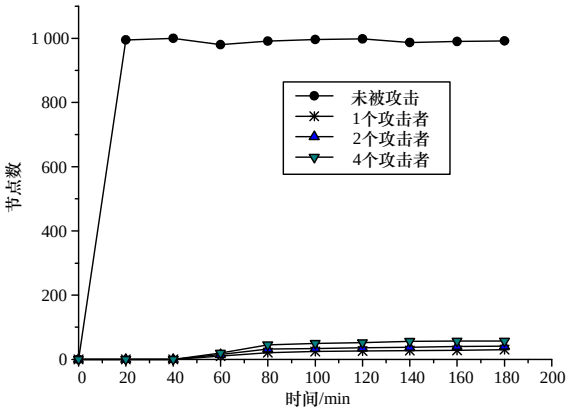

图 4　无防御系统时 Sybil 攻击的节点数与时间的关系

在 t=20 时，网络中有 1 000 个节点，并达到稳定状态。Sybil 节点于 t=40 分时开始进行攻击。本文实验结果显示，一个攻击者在 1 h 内可以控制系统近 10%的节点，而 4 个攻击者只需半小时就可以达到同样效果。图 5 给出了使用基于小世界模型防御机制情况下，P2P 系统中被控制节点的百分数。

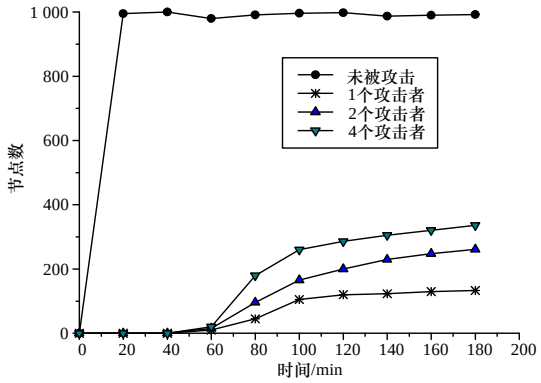

图 5　有防御系统时 Sybil 攻击的节点数目与时间的关系

如图 5 所示，一个攻击者在很长一段时间内只能控制系统中不到 2%的节点数，而 4 个攻击者对节点的控制能力不到 5%。因此，本文提出的基于小世界模型的防御机制能有效控制 Sybil 攻击的程度，降低 Sybil 攻击者控制的节点数在整个系统中的百分比。

参考文献

[1] Douceur J R. The Sybil Attack[C]//Proc. of the 1st International Workshop on Peer-to-Peer Systems. Cambridge, MA, USA: [s. n.], 2002: 251-260.

[2] Yu Haifeng, Kaminsky M, Gibbons P B, et al. SybilGuard: Defending Against Sybil Attacks via Social Networks[R]. Pisa, Italy: Intel Research Pittsburgh, Technical Report: IRP-TR-06-01, 2006.

[3] Kamvar S D, Schlosser M T, Garcia-Molina H. The Eigen Trust Algorithm for Reputation Management in P2P Networks[C]//Proc. of International World Wide Web Conference. Budapest, Hungary: [s. n.], 2003: 20-24.

[4] Faloutsos M, Faloutsos P, Faloutsos C. On Power-law Relationship of the Internet Topology[C]//Proceedings of the ACM SIGCOMM Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication. Cambridge, Massachusetts, USA: ACM Press, 1999: 251-262.

[5] Watts D J, Strogatz S H. Collective Dynamics of Small-world Networks[J]. Nature, 1998, 393(6684): 440-442.

[6] Adar E, Huberman B A. Free riding on Gnutella[Z]. Xerox PARC, 2000.

编辑　陈　晖

（上接第 102 页）

```
    else
        {
    //不是顺序语句，也不是 if 语句，那就证明是 switch 语句则调
//用分析 switch 语句的方法来分析语句，找到基路径
        }    }}
```

针对图 2，算法执行后可以得到基线路径：1→2→7→8→10→11→12。通过调用 FindOtherPath()这个函数可以旋转找到所有可能的基路径，再调用 eraseOverPath()去掉路径中重复的路径，得到的就是所要的基路径。

3　结束语

在白盒测试中，通过基路径覆盖技术可以缓解测试用例量过大与测试不足的矛盾。基路径覆盖技术在程序规模大的情况下更能体现出其优势。因此，采用基路径覆盖技术进行白盒测试能在保证测试充分性的前提下减少工作量，花较小的代价获得较高的软件质量。

参考文献

[1] Jorgensen P C. 软件测试[M]. 2 版. 韩　柯, 杜旭涛, 译. 北京: 机械工业出版社, 2003.

[2] Pressman R S. Software Engineering: A Practitioner’s Approach[M]. Beijing, China: Tsinghua University Press, 1996.

[3] 严蔚敏, 吴伟民. 数据结构(C 语言版)[M]. 北京: 清华大学出版社, 1996.

编辑　金胡考