P2P 环境下抗共谋攻击的自适应信任反馈聚合方法*

周 慎，祝跃飞
（国家数字交换系统工程技术研究中心，郑州 450002）

摘 要：针对信任管理有效防御模型共谋攻击的问题，提出了一种动态自适应的信任反馈聚合方法。通过引入二分网络投影对反馈聚合问题进行了合理转换，并根据反馈数据的集中程度自行调整参数，在无须人工干预的前提下，实现对信任反馈信息的过滤。仿真实验表明，该方法增强了信任管理模型对于共谋攻击的抵御能力，保证了模型在信任聚合阶段的鲁棒性。

关键词：共谋攻击；信任聚合；自适应
中图分类号：TP393　　**文献标志码**：A　　**文章编号**：1001-3695(2013)02-0494-03
doi:10.3969/j.issn.1001-3695.2013.02.048

Adaptive P2P trust aggregation method for resisting collusion attack

ZHOU Shen, ZHU Yue-fei
(*National Digital Switching System Engineering & Technological Research Center*, *Zhengzhou* 450002, *China*)

Abstract: Aiming at the lack of existing trust management model in resisting the collusion-attacks, this paper designed a dynamic adaptive trust aggregation method. According to the degree of concentration of the data, after problem-transformation with bipartite network projection, this method could automatically adjust the parameters and filter the data carrying trust feedback. The simulation results show that this method enhances the trust management model for the resilience of the collusion attack, to ensure the robustness of the model in the trust aggregation stage.

Key words: collusion attack; trust aggregation; self-adaptive

0 引言

近年来，P2P 技术凭借其灵活开放的组网模式、高度自治的体系结构，在文件共享、流媒体传输和即时通信等诸多领域均表现出令人期待的应用前景。但正是 P2P 网络“自愿参与、自主协同”的组织原则，为病毒、垃圾数据等提供了滋生的温床。已有研究表明[1~3]，建立有效的信任管理体系可以改善 P2P 网络的安全防护水平，但是各种针对信任管理体系本身的攻击行为制约了其进一步发展[4,5]。其中最为典型的是共谋攻击(collusion attack)，表现为若干个恶意节点达成协议，组成共谋团体，通过有组织地提交虚假信任反馈，控制某个目标节点的信任评价结果。同时，共谋节点间互为伪装、相互掩护，也提高了识别和抵御该类攻击的难度，当共谋团体足够庞大时，甚至能够颠覆整个信任管理机制。

针对 P2P 信任管理体系中共谋攻击防御难的问题，本文引入二分网络投影方法(bipartite network projection)[6]，提出了一种抗共谋攻击的信任聚合方法，通过扩充信任反馈信息的数据维度，增加能够体现反馈数据集中程度的参数标志，实现了反馈加权的动态自适应。仿真实验表明，该方法能够有效抵御针对信任管理体系本身的、有组织、有策略性的共谋攻击。

1 相关工作

由于 P2P 网络中的信任模型大多采用分布式应用结构，非协同的攻击手法作用有限且易于识别。相反，数量和组织的优势，使得大量共谋节点能够采取更加复杂隐蔽的攻击策略，产生更加严重的破坏效果。共谋攻击已经成为 P2P 信任管理所面临的重要威胁。现有的很多研究工作都对其进行了比较深入的研究和阐述，并提出了以下一些初步的解决方案：

EigenTrust[7] 信任模型提出了预置可信节点的方法，以抵御共谋团体的攻击，但是预先选定可信节点本身是个主观行为，缺乏必要的数据支撑；可信节点成为网络中的特权节点，一旦被攻陷，更易导致单点失效问题。

Lian 等人[8] 基于名为 Maze[9] 的 P2P 文件共享系统，深入研究了存在中心控制节点的网络环境下多个节点的共谋行为，归纳出几种常见的节点共谋模式(collusion pattern)，并提出了相应的检测方法[9]。但是这些方法并不适用于非集中式的 P2P 组织形式。该文献也证实了 EigenTrust 模型对于共谋攻击的抵抗能力较差。

XRep[10] 信任模型使用节点 IP 作为检测的依据，以识别共谋团体。即对节点 IP 进行聚类分析，属于同一聚类的节点被认为是来自同一个共谋团体。显然，这种方法对于 IP 地址欺骗(IP-spoofing)很脆弱，而且认为共谋节点在地理上一定相近本身就是欠妥的。

苗光胜等人[11] 引入模糊逻辑和语言变量，综合分析节点之间存在共谋行为的可能性，提出了基于模糊逻辑的共谋团体识别方法。但该方法引入的主观判断标准需要人为指定，尚不能做到动态自适应。

收稿日期：2012-07-12；**修回日期**：2012-08-23　　**基金项目**：郑州市科技创新团队项目(10CXTD150)

作者简介：周慎(1984-)，男，河南信阳人，硕士研究生，主要研究方向为信任管理、数据安全(kabinlosky@163.com)；祝跃飞(1962-)，男，浙江杭州人，教授，博导，博士，主要研究方向为密码学、信息安全、计算数论.

传统的研究方向侧重于共谋节点的筛选和共谋团体的划界。受制于 P2P 网络固有的关联动态性和拓扑复杂性,准确划定团体边界的难度较高,网络负载开销较大。本文着眼于共谋行为的固有特点,提出了无须判定节点所属阵营的协同防御办法,在实现所有参数动态自适应的同时,规避了传统方法所面临的效率与准确度的两难问题。

2　抗共谋攻击的信任反馈聚合

2.1　概念与定义

定义 1　信任评价。它是服务使用者(以下称 user)对服务提供者(以下称 provider)的服务质量所作的综合评价。User i 对 provider j 的信任评价记为 T_{ij}。

定义 2　信任反馈。它是 user 之间传递的消息,消息结构为形如(T_{ij},N_{ij})的二元组,其中 N_{ij} 表示 U_i 对 P_j 的反馈计数,初始值为 0,$N_{ij} \in \mathbb{Z}$ 。

定义 3　信誉。它是由信任反馈计算得到的,用于表征 provider 可信度的综合评价值。Provider i 的信誉记做 R_i,$R_i = f(T_{1i},T_{2i},T_{3i},\cdots)$,其中 f 是对应于信任管理方法的函数。

本文研究的重点是信任反馈的聚合方法,因此不对信任的表示和信任的更新作相关约束。但需要说明的是,此两者并不影响本文对所关注问题的研究结果:a)如何计算 T_{ij} 并不影响聚合算法的通用性;b)R_i 的更新发生于信任聚合和网络交互成功之后,处理对象是聚合算法的输出,与输入无相关性。

2.2　基于二分网络投影的问题转换

在任一时刻,P2P 网络中的节点可分为服务使用者和服务提供者两类,符合二分网络投影的应用要求[12]。如图 1 所示的拓扑结构,图中顶点表示网络节点,边表示通过交互产生信任评价。

将此拓扑向 user nodes 平面投影,结果如图 2 所示。图中各边表示信任反馈关系,对应的数字表示投影重叠次数。显然 U-P 关系已转换为 U-U 关系。

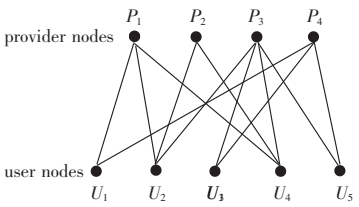

图 1　U-P 关系

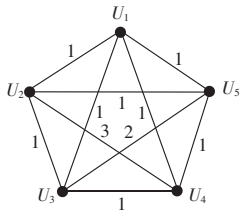

图 2　U-U 关系

以 U_2 与 P_3 交互为例,U_2 欲知 R_3,则信任聚合问题转换为确定 U_i 的信任反馈对 U_2 的影响程度问题,其中 $U_i \in \{U_p \mid T_{p3} \neq \text{null}\}$。

本文即通过这种转换,建立查询方(U_2)与反馈提供方(U_i)的直接联系,避免因信任反馈搜集机制的不同,降低聚合算法的通用性,同时也作为介绍算法时的背景约束。

2.3　抗共谋的自适应聚合算法

共谋攻击的主要目的是通过多节点的协同行为,控制目标节点信誉值。已有的研究成果表明[13],共谋成功需要两个前提条件:a)共谋节点间相互信任;b)共谋节点本身的信誉值不能过低。

本文提出一种抗共谋的自适应聚合算法,从共谋节点完成这两个前提的不同阶段入手,对共谋行为产生的信誉收益进行抑制:

a)获取信任反馈数据集

节点 U_i 按照网络在用的资源搜索协议发出对节点 P_j 信誉的查询请求,T_{pj}非空的节点响应此消息。

设 G_j 为 U_i 查询 R_j 所得信任反馈数据集,D_i 为 U-U 关系图中顶点 U_i 的维度,有

$$G_j = \{(T_{pj},N_{pj}) \mid 0 \leqslant p \leqslant D_i, T_{pj} \neq \text{null}\} \tag{1}$$

b)表征反馈数据的集中程度

以 N_{ij}为横坐标,对 G_j 中数据进行合并,取纵坐标为 Y_x,则

$$Y_x = \| G_j^x \| = \| \{(T_{pj},x) \mid (T_{pj},x) \in G_j\} \| \tag{2}$$

c)构造函数,对 T_{pj}加权

该函数应满足以下条件:(a)函数峰值点自适应于式(1)中 N_{ij}的数据集中度;(b)函数峰值自适应于式(2)中参数 x 的样本多数;(c)函数曲线与样本点拟合度最大化。

构造函数:

$$W(x) = \frac{1}{\sqrt{2\pi}\delta} e^{\frac{(x-\mu)^2}{2\delta^2}} \tag{3}$$

令 $n = \max(N_{ij})$。

将节点个数之比看做概率,计算 N 的均值,使高权重区间向样本多数偏移。

$$P_{ij} = \frac{Y_x}{\sum_{x=1}^{n} Y_x}$$

有 $\mu = \sum_{i=1}^{n} P_{ij}N_{ij}$。

以 $\min\left(\sqrt{\sum_i (W(x_i) - P_{ij})^2}\right)$的最优解为 δ。

d)根据 T_{pj}与函数 $W(x)$,计算:

$$R_j = \left(\sum_i^{G_j} T_{ij} \times W(Y_i)\right) / \sum_i^{G_j} W(Y_i) \tag{4}$$

以上是单个节点信誉值计算方法。实际网络中,多个节点的信誉值计算是并发的。节点间交互完毕后,还需要进行信任关系的更新。由于已经脱离信任聚合的范畴,所以本文对此不再深入探讨。

3　仿真与实验

为分析本文算法对共谋攻击的防御效果,构造一个运行 Chord 协议的 P2P 网络,通过分别关注单节点信任累积阶段和多节点协同攻击阶段的两组实验,对其进行仿真分析。

3.1　仿真环境

实验网络由 1 000 个节点组成,其中设置 1 个节点数目不定的共谋团体,文中将此个数比计为 k。普通用户的交互对象是在 Chord 路由协议框架下随机选定的,路径搜索深度不高于 6。而共谋节点与普通节点不同,为体现其隐蔽性与策略性,设置共谋节点按既定比率 c 与团体内节点互相提供共谋评价,同时以压低评价的方式对团体外的目标节点进行诋毁。实验所用参数如表 1 所示。

表 1　实验参数设置

参数	缺省值
节点总数(N)	1 000
共谋节点在网络中的比例(k)	0%-80%
共谋节点共谋评价的比例(c)	50%,100%
P2P 路由协议	Chord
最大路径长度	6
信任评价(T)	0-1

仿真基于 Peersim1.0.3,采用 Java 语言实现,硬件环境为 Core2、4 核、8 GB。

3.2 信誉累积实验

图 3 为共谋团体中某一节点累积信誉的过程($c=100\%$),从图中可以看出本文方法与传统平均值(average)方法的对比情况。实验初期,在节点尚未表现出共谋特质时,本文方法不会大幅抑制信誉的累积速度,因而不存在错误抑制正常节点的问题;随着共谋行为不断显现,本文方法能够有效缩减节点的共谋收益,达到对其实施抑制的目的。

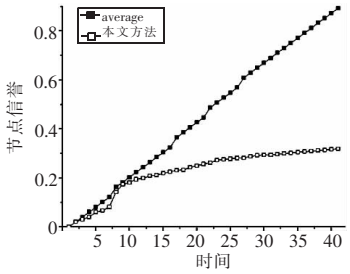

图 3 信誉的累积过程

图 4 所示为相同实验条件下,共谋团体的规模对实验结果的影响情况。可以看到,本文方法在 $k=20\%$ 时工作良好,经历 40 个轮询周期后,共谋节点的信誉值仍被控制在 0.31 的低位,即使共谋节点达到节点总数的 60% 时,本文方法仍然能够正常工作。

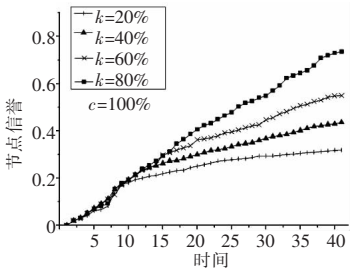

图 4 团体规模与信誉累积的关系

值得注意的是,当共谋团体的规模达到总节点数的 80% 后,算法产生明显抑制效用的时间会有延后,抑制共谋收益的实际效果受到影响,如表 2 所示。后续工作需要重点考虑这一问题。

表 2 信誉累积实验数据

团体规模	抑制的时机	节点信誉(40 周期)
$k=20\%$	11 周期	0.31
$k=40\%$	13 周期	0.43
$k=60\%$	15 周期	0.55
$k=80\%$	21 周期	0.73

3.3 协同攻击实验

模拟共谋节点协同诋毁团体外某一节点的信誉,采样点为遭受攻击的正常节点,用于攻击的消息比例为 50%,实验结果如图 5 所示。可以看出,本文方法对恶意诋毁行为的抑制作用明显,节点信誉的下降速率本身一直处于衰减状态。持续抑制的结果是共谋节点所造成的负收益比例不断减少,节点本身的良好行为所引发的正收益比例不断增加,经过一段缓冲期后,使得节点信誉开始恢复。实验中,节点信誉下降速度首次出现平滑趋势的时间较早(5 ~ 7 周期),随后第二次出现由快速下降向平滑的转变(20 ~ 23 周期),这是由于路由寻径已经遍历了所有实验节点,信任反馈数据的方差出现暂时的减小所致。

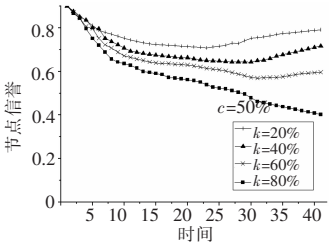

图 5 不同团体规模下的协同诋毁情况

4 结束语

本文以二分网络投影方法进行问题转换,提出一种基于数据集中度的节点信誉值的求解方法。通过增加信任反馈数据的维度,分析反馈数据的集中度,以自适应于样本多数的动态加权模式,对反馈结果进行了整形和过滤处理。仿真结果表明,此方法提升了信任管理模型在信任聚合阶段的抗共谋攻击能力。

参考文献:

[1] BLAZE M, FEIGENBAUM J, LACY J. Decentralized trust management[C]//Proc of IEEE Symposiums on Security and Privacy. Washington DC: IEEE Computer Society, 1996: 164-173.

[2] JØSANG A, HAYWARD R, POPE S. Trust network analysis with subjective logic[C]//Proc of the 29th Australasian Computer Science Conference. Hobart: Australian Computer Society, 2006: 85-94.

[3] MANCHALA D W. Trust metrics, models and protocols for electronic commerce transactions[C]//Proc of the 18th International Conference on Distributed Computing Systems. [S. l.]: IEEE Computer Society, 1998: 312-321.

[4] MEKOUAR L, IRAQI Y, BOUTABA R. Peer-to-peer's most wanted: malicious peers[J]. Computer Networks, 2006, 50(4): 545-562.

[5] SINGH A, CASTRO A, DRUSCHEL P, *et al*. Defending against eclipse attacks on overlay networks[C]//Proc of the 11th Workshop on ACM SIGOPS European Workshop. New York: ACM Press, 2004.

[6] ZHOU Tao, REN Jie, MEDO M, *et al*. Bipartite network projection and personal recommendation[J]. Physical Review E, 2007, 76(4): 104-115.

[7] KAMVAR S D, SCHLOSSER M T, GARCIA-MOLINA H. The eigentrust algorithm for reputation management in P2P networks[C]//Proc of the 12th International Conference on World Wide Web. New York: ACM Press, 2003: 640-651.

[8] LIAN Qiao, ZHANG Zheng, YANG Mao, *et al*. An empirical study of collusion behavior in the maze P2P file-sharing system[C]//Proc of the 27th International Conference on Distributed Computing Systems. Washington DC: IEEE Computer Society, 2007: 56-69.

[9] Maze[EB/OL]. [2012-07-10]. http://www.tianwang.com/.

[10] DAMIANI E, De VIMERCATI C D, PARABOSCHI S, *et al*. A reputation-based approach for choosing reliable resources in peer-to-peer networks[C]//Proc of the 9th ACM Conference on Computer and Communication Security. New York: ACM Press, 2002: 207-216.

[11] 苗光胜,冯登国,苏璞睿. P2P 信任模型中基于模糊逻辑的共谋团体识别方法[J]. 计算机研究与发展,2011,48(12):2187-2200.

[12] 吴亚晶,张鹏,狄增如,等. 二分网络研究[J]. 复杂系统与复杂性科学,2010,7(1):1-12.

[13] HOFFMAN K, ZAGE D, NITA-ROTARU C. A survey of attack and defense techniques for reputation systems[J]. ACM Computing Surveys, 2009, 42(1): 1-31.