Peer-to-Peer 信任模型中的恶意行为分析

余智华[1,2]
YU Zhi-hua[1,2]

1.中国科学院 计算技术研究所 智能软件部,北京 100080
2.中国科学院 研究生院,北京 100039
1.Software Division,Institute of Computing Technology,Chinese Academy of Sciences,Beijing 100080,China
2.Graduate School of the Chinese Academy of Sciences,Beijing 100039,China

YU Zhi-hua.Analysis of malicious behaviors in peer-to-peer trust model.Computer Engineering and Applications,2007,43(13):18-21.

Abstract: Through lack of strict identification and trust mechanisms,there are a lot of malicious peers in current P2P network. This paper introduces several typical trust models and malicious behavior patterns,then analyzes the characters of these trust models and their resistance to difference malicious behaviors by simulations.It shows that trust models based on feedback can effectively protect the network from many kinds of malicious behaviors.
Key words: Peer-to-peer;trust model;malicious behavior

摘 要:在目前已广泛应用的 P2P 网络中,由于缺乏严格的身份验证和信任机制,存在着许多欺诈等恶意行为,系统的有效性和可用性难以保证。论文总结了 P2P 网络中的信任模型和恶意行为模式,并针对不同的恶意行为分别对几种典型的 P2P 信任模型进行模拟试验,对比分析了 P2P 信任模型的特点和对各种恶意行为的抑制作用。结果表明基于反馈的信任模型能够有效地抵御多种恶意节点的破坏,提高网络服务的有效性。
关键词:对等网络;信任模型;恶意行为
文章编号:1002-8331(2007)13-0018-04 **文献标识码**:A **中图分类号**:TP393

1 引言

Peer-to-Peer(P2P)技术是近年来兴起的一种新型网络模型,其区别于传统 Client/Server 模型或 Browser/Server 模型的最显著特点是整个网络不存在中心节点(或中心服务器)、其中的每一个节点(Peer)大都同时具有信息消费者、信息提供者和信息通讯等三方面的功能。P2P 是一种完全分布式网络,其中的参与者共享他们所拥有的一部分硬件资源,通过网络提供服务和内容,能被其他 Peer 直接访问。

随着 P2P 应用的推广,P2P 网络中的安全问题也逐渐暴露出来。P2P 一般具有自治性、动态性等特性,节点的进入和退出往往不受任何限制。这些特性在提供了很大便利的同时,也在网络管理和安全保障方面带来了诸如知识产权保护、信息泄漏、病毒传播、恶意欺骗等一些问题。由于缺乏有效的身份验证和信任机制,P2P 网络中存在着大量的欺诈等恶意行为,这些行为大大降低了系统整体的可靠性和可用性。以目前互联网上应用广泛的文件共享应用为例,25%的文件是伪造文件(faked files)。在商业应用中,随着 C2C 电子商务的发展,信用问题尤为突出,美国联邦贸易委员会(FTC)2002 年收到了 5.1 万份相关投诉,总计 3 700 万美元的损失,其中 85%发生在 eBay 网站。因此,建立起有效的信任机制就显得至关重要。当前很多在线交易系统都采用基于反馈和推荐的等级评价或类似系统,如 eBay、Yahoo!、Auction 等,但通常比较简单,缺乏对信息的充分分析和对欺骗等恶意行为的抑制。

本文首先介绍了当前在信任模型方面的研究以及几种典型的 P2P 信任模型,并对 P2P 网络中存在的恶意行为的模式进行了归纳,最后通过模拟试验对比分析各种模型的优缺点和抵御恶意行为的能力。

2 信任模型研究

2.1 信任模型分类

为抑制网络中的恶意行为、提高节点和服务的可靠性,近年来国内外在信任机制方面进行了大量的研究工作,提出多种信任模型。目前提出的信任模型大致可以归为以下几类:

(1)基于可信第三方的信任模型采用 PKI 等技术,在网络中建立集中式认证服务器或少数领袖节点(leader peers)来监督整个网络的运行。这类模型依赖于少量中心节点,无法避免可扩展性、单点失效等方面的一系列问题。

(2)数据签名采用类似投票的机制评判信息的可信度,而不考虑节点是否可信。这种方法仅限于数据共享应用,而且无法抑制集体欺诈行为。

基金项目:国家高技术研究发展计划(863)(the National High-Tech Research and Development Plan of China under Grant No.2005AA142110)。
作者简介:余智华(1973-),男,副研究员,在读博士。

(3)基于群体信息共享的反馈和评价推荐机制搜集节点的历史活动信息和反馈作为建立信任的基础,经过一定的存活期后其获得的信任度逐步收敛到其真实值。

基于反馈信息的信任模型具有很好的应用灵活性、健壮性和较低的成本,而准确度较低和需要收敛时间等问题也可以通过适当的机制和算法加以弥补,适合 P2P 模式分散化、动态可伸缩的特点,是 P2P 应用中信任技术的主要发展方向,下文将对该类信任模型做进一步分析。

2.2 基于本地或局部推荐的信任模型

通过节点的历史活动信息来评价节点的可信度,最简单和直观的方法是在本地记录本节点与目标节点之间的所有交易,统计成功率,从而计算与其进行下一个交易成功的先验概率。

令 S_{ij} 为节点 i 和节点 j 交易的成功次数,F_{ij} 为两者交易的失败次数,$I_{ij}=S_{ij}+F_{ij}$ 为交易总次数。则节点 i 对节点 j 的本地信任度为

$$P_{i,j}=\frac{S_{i,j}}{I_{i,j}} \tag{1}$$

基于本地信任的信任模型最大的好处就是不会受虚假反馈的影响,因为所有反馈信息都是本地产生和累计的。因此在不同的应用环境和恶意行为下表现比较稳定。但该模型存在的最大问题是反馈信息的积累比较慢,必须要两个节点间反复进行交易才能获取足够的信息以保证先验概率计算的准确性。在稍大规模的 P2P 网络中,则很难获得对网络中所有节点的完整的信任信息。

为加快反馈信息的积累,减少建立信任的延迟,可以通过询问有限的其他节点、共享反馈和评价信息以获取某个节点的可信度。在这类基于局部推荐的信任模型中,往往采取简单的局部广播的手段。在 Cornelli 提出的针对 Gnutella 的信任模型 P2PRep 中采用的就是这种方法。基于局部推荐的信任模型通过限制反馈和评价信息的共享范围,使得由此带来的通讯开销不致过大,但其获取的节点可信度也往往是局部的和片面的。

2.3 全局信任模型

全局信任模型则是在整个网络范围内共享反馈和评价信息,对每个节点建立起全局的信用评估,如 Stanford 的 eigenRep 等。

要让整个网络中的节点共享全局信任度,必须解决信用信息的检索定位问题,并避免信用信息被篡改。一般来说,可以采用分布式 Hash 表(DHT)来组织,如 Chord、CAN 等,利用单向 Hash 函数将全局信用信息节点存放到另一个可定位的节点上。

Li Xiong 等提出了一个泛化的全局信任度计算模型:

$$T(u)=\alpha*\sum_{i=1}^{I(u)}S(u,i)*Cr(p(u,i))*TF(u,i)+\beta*CF(u) \tag{2}$$

其中 $I(u)$ 表示节点 u 与其它节点间的事务(Transaction)总数;$S(u,i)$ 表示 u 第 i 次事务获得的满意度反馈(已规格化到[0,1]);$p(u,i)$ 表示节点 u 第 i 次事务的对象节点;$Cr(v)$ 表示 v 提供的评价的可信性;$TF(u,i)$ 为 u 第 i 次事务的调整因子(Transaction Context Factor),可以是传输内容长度等,根据事务的重要程度赋以不同的权重;$CF(u)$ 表示 u 的全局调整因子(Community Context Factor);α、β 为用于规格化的权重因子。

从公式(2)可以看出,全局信任度 $T(u)$ 由两部分组成:前一部分由反馈计算出的信任值赋以权重 α;后一部分则可根据节点的其它情况进行适当的奖励或惩罚,还可用于确定信任度的初始值。

在公式(2)中,Cr 函数可以有多种选择,常规的做法是假设正常节点提供的反馈信息是真实的,通过迭代逐步排除恶意节点发布的虚假信息,最终可信度会收敛到与其历史表现相一致。即将 Cr 定义成 T 的函数,从而得到以下公式,称之为 TVM(Trust Value Measure)算法:

$$T_{TVM}(u)=\sum_{i=1}^{I(u)}S(u,i)*\frac{T(p(u,i))}{\sum_{j=1}^{I(u)}T(p(u,j))} \tag{3}$$

2.4 在信任模型中引入聚类思想

常言道“物以类聚,人以群分”,拥有共同兴趣的人相互间往往能够具有更高的可信度。在网络信任模型中也可以引入这种聚类的思想,使得具有共同兴趣的节点其反馈的可靠性得以提高。具体做法就是将公式(2)中的 Cr 函数定义为节点相似度 $Sim(v,w)$ 的函数,得到 PSM(Personalized Similarity Measure)算法:

$$T_{PSM}(u,w)=\sum_{i=1}^{I(u)}S(u,i)*\frac{Sim(p(u,i),w)}{\sum_{j=1}^{I(u)}Sim(p(u,i),w)} \tag{4}$$

$$Sim(v,w)=1-\sqrt{\frac{\sum_{x\in IJS(v,w)}\left(\frac{\sum_{i=1}^{I(x,u)}S(x,i)}{I(x,v)}-\frac{\sum_{i=1}^{I(x,w)}S(x,i)}{I(x,w)}\right)}{|IJS(v,w)|}} \tag{5}$$

其中:$IJS(v,w)$ 表示与节点 v 和 w 都有过交易和反馈的节点集合。

计算节点相似度时要大量地获取其它节点的反馈信息,通讯开销和延迟是非常大的,在具体实现时可以采用缓存机制将节点相似度缓存到本地从而提高系统性能。此模型存在的另一个问题是存储节点两两间的相似度所产生的开销,一个优化的思路是对节点聚类,每个节点归入一个类,类中的信任信息保持一致、相互共享。

目前来看,将聚类思想引入网络信任模型具有一些独特的优点,但算法还不成熟,还有许多工作要做。

2.5 应用实例:eBay 的信用机制

作为世界上最大的网上交易平台,eBay 公司创办者特别强调信任问题,将其网上拍卖业务称为是建立在相互信任基础上的电子商务的一次成功实践。

eBay 所采取的信用机制非常简单:

(1)集中管理所有注册用户的信用信息。

(2)每个用户都可以根据交易中的表现给对方评分,分为好、中、差 3 级。

(3)所有其他用户对某个用户的评分累加起来就得到该用户的信用度。

(4)一个用户只能对另一个用户的信用度造成 1 分的影响,即使给出多次好评也只会加 1 分。

该信用机制本质上是一个基于集中管理的全局信任模型,算法简单、实现效率很高。但其未解决评价信息可靠性的判别问题,只是通过限制两个用户间评分的累计来减少不可靠评价带来的影响。

3 恶意行为分析

不同于由意外或故障引起的破坏,网络中的恶意行为往往

有较强的目的性,其行为模式也具有主观性和不确定性。因此,试图对攻击者的行为进行真实、准确的定义是一件非常困难的事情。为细致地分析信任模型的效果和优缺点,将可能出现的恶意行大致规纳为以下几大类。

3.1 欺骗

这一类恶意节点最为常见,以在交易过程中欺骗对方节点为目的,或是在文件共享网络中提供假冒的文件下载服务。这类节点是一般的信任模型中首先要识别的目标。

3.2 合谋

合谋(collusive)是指多个恶意节点之间相互串通,互相给予对方很高的反馈和评价,以此来抬高双方的可信度。由于在P2P网络中不限制节点的加入和退出,攻击者可以建立大量的影子节点来协同欺骗,而且可以在合谋节点间进行假交易而迅速抬高恶意节点的可信度。

要抑制此类合谋欺骗行为,一般的方法是对节点间的评价反馈加以一定的限制,如限制评价值的大小和反馈次数,甚至象eBay的信任模型那样限定一对节点间有效的评分只有1分。

3.3 诋毁

诋毁行为是指恶意节点在与正常节点交易或请求下载服务时,即使下载成功也反馈负面评价,以此来压低对方的可信度、破坏正常节点的服务。

要防止诋毁行为的危害,就必须对节点的反馈评价进行有效地甄别。如前文所述,在大部分信任模型中都是基于节点的可信度来决定其反馈的可靠性,如果恶意节点在提供正常服务的伪装下对其它节点进行诋毁,则很难将其识别出来。2.2节中介绍的基于反馈相似度的甄别方法另辟蹊径,在识别诋毁行为时无需考虑其所提供服务的可信度,是一个有益的尝试。

3.4 其它形式的恶意行为

除了上面归纳的几类主要行为模式外,攻击者还可能采取冒名、潜伏、反复重入等方式。

3.4.1 冒名

恶意节点冒充正常节点提供假的文件下载或交易信息,或是发布假冒的评价反馈。在通讯过程中增加对信息的确认、签名或其它身份认证手段即可较好地防止冒名攻击的破坏。

3.4.2 潜伏

恶意节点在进行欺骗时可以有一定的潜伏期,间歇性地提供假信息和服务。这种行为比单纯的欺骗更具有隐蔽性,可能使恶意节点的欺骗行为被正常服务获得的正反馈所掩盖,从而维持较高的信任度。可采用多重窗口的机制,使得发生欺骗行为后节点的信任度迅速下降,而在获得正反馈时信任度的回升速度则慢得多,从而使恶意节点难以维持高的信任度。

3.4.3 重入

在攻击者实施恶意行为导致节点信任度较低时,可以从P2P网络中退出再以一个新的身份进入,即可摆脱原先节点的低信任度。对于类似eBay的信用机制,采用重入方式可以避开其"一人一票"的限制,进而进行合谋、诋毁等攻击行为,因而具有更大的危害性。这种恶意行为可以通过适当设置节点信任度的初始值、增加退出/再进入的开销和代价、乃至用户身份验证等手段加以遏制。

3.4.4 Free Riding

在P2P网络中除恶意行为外还有一类"搭便车者"(free-riders),只希望使用网络的资源而自己并不想提供资源。这些行为虽然并不会给网络带来多大破坏,但它违背了P2P"人人为我,我为人人"的精神,受到普遍的批评。而且大量free-riders的存在也影响了资源的平衡,降低了网络的整体性能。在信任模型中,节点提供的有效服务或交易越多,收到的正面反馈评价越多,其信任度就越高。据此可以很容易地确定节点对网络的贡献大小,制定相应的奖惩措施。

4 信任模型的模拟与分析

为了分析各种信任模型对恶意行为的抑制作用,针对几种主要的恶意行为模式进行了模拟试验。模拟试验设定的场景如下:

(1)在包含 n 个节点的P2P网络中,每个节点都可以向其它节点提供服务。

(2)进行 m 轮模拟,每一轮中每个节点 P 发出一次服务请求,从随机产生的 q 个响应节点中选择信任度最高的节点请求服务,服务结束后节点 P 根据服务成功与否对服务提供节点进行评价,评价值介于[0,1]区间。

(3)对信任模型的评价标准是整个网络中服务的成功率。服务成功率越高,说明模型更能够抑制恶意行为,避免其对正常业务的影响。

模拟试验的参数设置如表1。

表1 模拟参数

参数	说明	默认值
n	节点总数	100
m	模拟轮数	100
q	响应节点数	5
k	恶意节点占总节点数的比例	40%

在对比试验中选择了本地信任(Local)模型、eBay模型、TVM、PSM 4个比较有代表性的信任模型。Local模型只相信自己的经验,不需要其它节点的信用、评价信息,因而比较稳定、不易受恶意行为的影响,可以用来作为对比的参照。

4.1 欺骗行为模拟

假设网络中的恶意节点都是单纯的欺骗节点,即服务是虚假的,而反馈是真实的。随着恶意节点比例的增加,各信任模型中服务的成功率变化如图1所示。

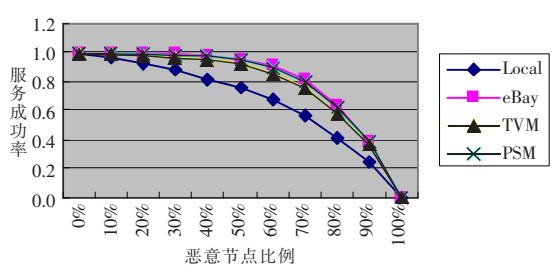

图1 普通欺骗行为下的服务成功率

由模拟结果可以看出,全局信任模型明显优于本地信任模型。由于无论正常节点还是恶意节点的反馈信息都是真实的,就避开了反馈可靠性甄别的问题,三种全局信任模型的效果相近。

4.2 合谋行为模拟

在此项模拟中,假设所有恶意节点都串通到一起,相互之间给予正反馈。随着恶意节点比例的增加,各信任模型中服务的成功率变化如图2所示。

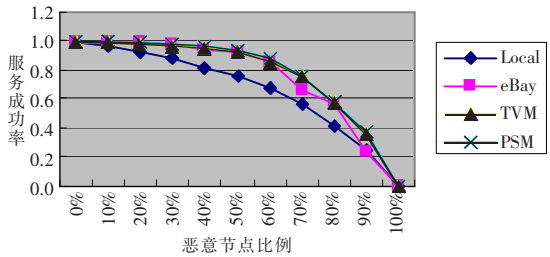

图 2 合谋欺骗行为下的服务成功率

将图 2 中结果与图 1 对比可以发现，两者的曲线非常相近,可见合谋行为造成的服务成功率下降幅度不大。只有 eBay 算法因无法甄别反馈评价信息的真假,受到的影响比较明显。

4.3 诋毁行为模拟

考虑到诋毁行为往往都是多个节点协同进行的,否则不会产生明显的效果,由此本项试验中,设定恶意节点都是合谋的并对所有正常节点发出负反馈进行诋毁。模拟结果如图 3 所示。

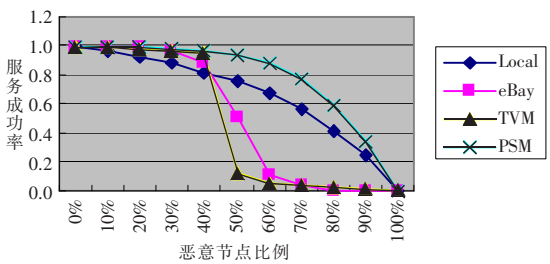

图 3 合谋诋毁行为下的服务成功率

结果中最明显的变化是：当恶意节点比例达到 50%左右时,eBay 和 TVM 模型的状态发生了一次跃变，由 90%以上迅速下跌到 20%以下,可以说系统已经基本失效了。究其原因,当恶意节点占据了节点中的大部分的时候,诋毁的反馈评价已经压倒了正常反馈,导致信任模型中的信任度和反馈可靠度同时被翻转。

在此项试验中,PSM 模型表现得一枝独秀，一直到恶意节点占很高比例时仍能保持不错的服务成功率。进一步分析其信任度和相似度发现,网络中的节点实际上被划分成正常和恶意节点两大类,同类的节点间信任度很高,而对另一类的节点信任度则很低。因而正常节点在请求服务和计算信任度时依赖于其它的正常节点,基本不受恶意节点的影响。

4.4 信任模型收敛速度分析

在上述模拟试验中,节点间能够进行充分的交互,获得足够的反馈使得计算出的节点信任度能够收敛。但在大多数实际应用中,节点往往需要在不能获得足够反馈信息的情况下判定节点是否可信。为此,针对信任模型的收敛速度进行了试验:恶意节点为单纯的欺骗节点，节点总数 n=200，总模拟轮数 m=200,模拟结果如图 4 所示。

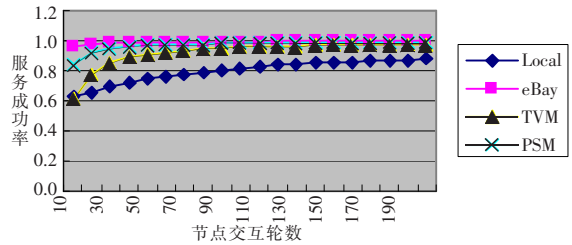

图 4 服务成功率随节点交互轮数的变化

随着节点交互的增加，全局信任模型能够较快地收敛,在每个节点与网络中约 20%的节点进行过交易后,已经可以通过评价信息的共享获得所有节点的较准确的信任度。而 Local 模型因为缺乏信息共享,需要节点进行大量的交互后才能积累足够的反馈信息。

5 结论

表 2 信任模型模拟分析结果

信任模型	说明	收敛	抗恶意行为		
			欺骗	合谋	诋毁
Local	本地信任	慢	较好	较好	较好
EBay	全局信任,集中式管理	快	好	好	较差
TVM	全局信任	较快	好	好	较差
PSM	全局信任	快	好	好	好

本文分析了几类典型的信任模型,并模拟了多种主要的恶意行为对模型的影响,试验结果如表 2。不难看出:本地信任模型和全局信任模型相比较,全局信任模型在收敛速度以及抑制大部分恶意行为方面占有优势，而本地信任模型则相对稳定、不受合谋诋毁等影响。带有聚类思想的信任模型表现出对大规模合谋恶意行为的较好抵抗力。总体来看,全局信任模型能够大范围地共享信用信息,具有更好的效果。

(收稿日期:2007 年 1 月)

参考文献:

[1] 王晓燕.CtoC 电子商务中的信任问题:一个进化博弈分析模型[J].商业研究,2005(314):179-181.

[2] Sulin B,Andrew B W,Han Z.The dynamics of the electronic market:an evolutionary game approach[J].Information System Frontiers,2000,2:31-40.

[3] 窦文,王怀民,贾焰,等.构造基于推荐的 Peer-to-Peer 环境下的 Trust 模型[J].软件学报,2004,15(4).

[4] Sepandar D K,Mario T S,Hector G.The eigenTrust algorithm for reputation management in P2P networks[C]//12th International World Wide Web Conference.Budapest:ACM Press,2003:123-134.

[5] Li Xiong,Ling Liu.PeerTrust:supporting reputation-based trust for peer-to-peer electronic communities[J].IEEE Transactions on Knowledge and Data Engineering,2004,16(7).

(上接 13 页)

[5] Abdul-Rahman A,Hailes S.Supporting trust in virtual communities[C]//Proceedings of the Hawaii International Conference on System Sciences,Maui,Hawaii,4-7 January 2000.

[6] Kamvar S D,Schlosser M T.EigenRep:reputation management in P2P networks[C]//Lawrence S.Proc of the 12th Int'l World Wide Web Conf.Budapest:ACM Press,2003:123-134.

[7] Wang Y.Bayesian network-based trust model in Peer-to-Peer networks[C]//Workshop on "Deception,Fraud and Trust in Agent Societies" at the Autonomous Agents and Multi Agent Systems,2003.