软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software,2011,22(10):2346−2357 [doi: 10.3724/SP.J.1001.2011.03901]

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel/Fax: +86-10-62562563

一种 P2P 文件共享网络高精度自适应声誉机制*

王 淼[1,2+], 陶 飞[1,2], 张玉军[1], 李国杰[1]

[1](中国科学院 计算技术研究所,北京 100190)

[2](中国科学院 研究生院,北京 100049)

Accurate and Adaptive Reputation Mechanism for P2P File Sharing Network

WANG Miao[1,2+], TAO Fei[1,2], ZHANG Yu-Jun[1], LI Guo-Jie[1]

[1](Institute of Computing Technology, The Chinese Academy of Sciences, Beijing 100190, China)

[2](Graduate University, The Chinese Academy of Sciences, Beijing 100049, China)

+ Corresponding author: E-mail: wangm@ict.ac.cn, http://www.ict.ac.cn

Wang M, Tao F, Zhang YJ, Li GJ. Accurate and adaptive reputation mechanism for P2P file sharing network. *Journal of Software*, 2011,22(10):2346−2357. http://www.jos.org.cn/1000-9825/3901.htm

Abstract: Current reputation schemes, employed by an existing P2P file sharing network are faced with many threats, such as aggregate feedback, conspiracy, and fake transactions, which have affected the performance of whole system. To protect the P2P file sharing network, this paper proposes an accurate and adaptive reputation mechanism (AARep). In the paper, after an elegant analysis, it is concluded that besides the transaction evaluation, the relative and underlying transaction information plays an important role in the reputation system. This paper makes two important contributions. First, trust value computation is enhanced by adopting a decay function to display the importance of every transaction in sequence order. This weeds out suspected nodes that are less similar and utilizes the confident factors of reflect reliability in all observed values. Second, to make the confidence factors work, a simple transaction validation protocol is developed. Experimental results show that AARep can significantly eliminate or minimize the effects of a verity of attacks and improve the overall performance effectively.

Key words: file sharing; P2P network; similarity; transaction validation protocol; trust value

摘 要: P2P 文件共享网络的信任评价机制正面临着各种恶意攻击,例如聚集反馈、合谋欺骗和虚假交易,严重影响了整个系统的性能.为了保护 P2P 文件共享网络,提出了一种高精度自适应声誉机制,简称 AARep 机制(accurate and adaptive reputation mechanism).经过分析发现,在信任评价系统中,除了对交易的评价,相关的交易信息也起着重要的作用.其贡献在于:1) 通过以下手段,增加了信任值计算的精度:引入交易衰减函数,根据交易次序区分交易的重要程度;过滤掉具有低相似度的可疑节点;利用置信因子来反映观测值的可靠性;2) 为了使置信因子起效,提出了一种简单的交易验证协议.实验结果表明,AARep 能够显著消除或者减少各种攻击的威胁,提高系统性能.

关键词: 文件共享;P2P 网络;相似度;交易验证协议;信任值

* 基金项目: 国家自然科学基金(60803139); 国家重点基础研究发展计划(973)(2007CB310702); 国家科技支撑计划(2008BAH37B02)

收稿时间: 2009-11-06; 修改时间: 2010-03-05; 定稿时间: 2010-06-28

中图法分类号: TP393 文献标识码: A

在 P2P 网络中,每个节点(peer)地位相等,同时扮演着资源提供者(服务器)和资源请求者(客户端)的角色,消除了传统 C/S 模式下服务器单点失效和可扩展性差等问题.目前,大多数 P2P 网络在设计时都假定所有节点按照协议正确工作,即资源提供者一直提供真实且高质量的内容和服务.但由于 P2P 结构的开放性、节点的匿名、自治性等特征,这个假定是不现实的,恶意用户可能散布虚假或者伪劣的内容和服务.例如:对 KaZaA[1]的研究表明,有超过 50%的音频文件是被污染的(polluted)[2];在 P2P 文件共享系统中,VBS.Gnutella 蠕虫病毒[3]的流行等.研究表明,建立有效的信任评价机制、标识出可靠节点,是提高 P2P 网络整体可用性和可靠服务质量的有效途径[4].

研究人员已经提出了一些用于 P2P 网络的信任评价机制[5−13].大多数自组织 P2P 系统在信任值计算时,为了使最后的结果更客观,通常用一个表示反馈源可靠性的因子,例如服务声誉或者与请求者的相似度,对提交的反馈进行加权,使得更合作的节点能够得到更大的权重.但是,这些方法存在一些问题,使得信任机制对某些攻击不够敏感.例如:很多机制在信任值计算时,并不区分经验丰富的推荐者和无经验的推荐者,而且把暗示着不可信的低相似度节点的推荐也包含也内;在相似度计算时,认为所有的评价都是相同的,无论其背后进行了多少次交易;服务质量突然下降时,信任值并不会迅速发生改变.

目前,很少有研究关注交易的真实性.虽然也有一些机制[14]要求交易双方交换信息、核实身份,但是这些机制是专为电子商务或电子邮件类应用设计的.对 P2P 文件共享来说,这个要求过于严格,特别是当我们只关心两个节点之间是否真的发生过交易的时候.

为了解决这些问题,本文提出了一种 P2P 文件共享网络高精度自适应声誉机制(an accurate and adaptive reputation mechanism for P2P file sharing network,简称 AARep 机制).我们首先分析了信任值的推理过程,强调交易信息的重要性.然后,AARep 在信任值的计算时包括了更多的交易信息,改善了信任值计算的准确性和适应能力.最后,为了验证交易的真实性,提出了一种轻量级的交易验证协议.该协议不需要任何可信的第三方参与,而且与用于商业领域的协议相比,更加简单.

1 相关工作

根据历史信息共享范围的不同,P2P 网络的信任评价机制大致分为如下 3 类:

1) 基于中心节点的声誉机制.这类机制中存在少数中心节点负责监督系统的运行,并定期通告违规的节点.

这类系统如 eBay[5],Epinions[6]等往往是中心依赖的,具有可扩展性差、单点失效等问题.

2) 基于全局信息的声誉机制.这类机制通过邻居节点间交易满意度的迭代,获取节点的全局信任值.

Kamvar 等人[7]基于信任的传递性,提出了 EigenTrust 机制.在信任值计算时,EigenTrust 以节点的全局信任值本身作为推荐的权重,假设服务性能与推荐可信度之间存在相关性,即服务好的节点也会提供真实的反馈;提供坏服务的节点汇报的反馈也是不可信的.但这个假设不总是成立的.在无恶意行为的网络中,该机制计算得到的信任值可以较好地反映节点的真实行为.但该机制具有较高的通信代价(每次交易都会导致全网络的迭代),而且其安全(及收敛)保证依赖于网络中预先存在若干具有较高声誉的中心服务器节点,这在 P2P 环境下很难实现.窦文等人[8]在迭代收敛性和安全性方面对 EigenTrust 进行了改进,但改进后的模型仍然存在效率问题,且其安全性是通过引入额外的认证机制和惩罚措施来实现的.PowerTrust[9]通过一种分布式等级机制,动态地选择出一些高声誉的节点.利用这些节点,使用前向随机游走策略,PowerTrust 提高了全局信任值的准确性和聚合速度.基于全局信息的声誉机制的共同缺点是忽略了信任值的私人化特征,对于某个特定的节点,其他节点对它的信任值都是相同的.

3) 基于局部信息的声誉机制.这类机制通过查询有限的其他节点,计算出节点的相对信任值.

Xiong 等人[10]提出的 PeerTrust 机制综合考虑了影响信任值的多个因素:交易的评价、资源提供者的交易

次数、反馈源的可信度、交易上下文和社区上下文,并提供了一种纯分布式的信任值计算方法.PeerTrust 基于个人相似度来计算信任值,即对那些曾经就一些公共交易节点给出过相似评价的推荐者,赋予更高的推荐权重.Sorcery[11],SFTrust[12]等机制也区分了服务性能和推荐信任.但是,以上算法仅考虑了交易评价,忽略了影响评价质量的交易次数、交易次序等因素.

在交易真实性方面,PeerTrust 规定了如何安全地提交和传输反馈(没有涉及交易验证).TrustGuard[13]通过交换交易证据来验证交易的真实性,但是,该方法需要一个可信的第三方,强制交易双方在每次交易过程中交换产品和收据.

2 问题描述

信任评价机制正面临着各种恶意攻击.攻击者可能更策略地、更隐蔽地发送误导性评价.

2.1 聚集反馈(aggregate feedback)

一个典型的攻击方法是,节点首先通过一段时间的成功交互获得较高的信任值,然后滥用自己的信任值去欺骗其他节点.聚集反馈攻击能够使节点保持一定的声誉,不断从系统中获得好处.如果只考察节点长期、完美的交易历史,就不容易察觉这种攻击,导致一些不幸的节点被欺骗.信任评价机制的责任是防止恶意文件的蔓延,这就要求它准确、及时地反映节点的状态.

2.2 合谋欺骗(conspiracy cheat)

合谋欺骗的节点形成恶意团伙,对团伙内的节点提交高推荐,对团伙外的节点进行诋毁.合谋欺骗团伙可以在较短时间内产生严重的攻击效果.如果合谋团伙非常大,那么它甚至会摧毁声誉系统.幸运的是,实际的 P2P 网络中大多数节点是诚实的,因此能够帮助声誉系统识别不可信节点及其反馈.

2.3 虚假交易(fake transactions)

恶意节点可能会伪造虚假交易.或者,在一些对用户注册不进行限制或限制较少的 P2P 网络中,节点可以通过被称为 Sybil 的攻击方式创建大量的虚假身份来破坏声誉系统.攻击者可以伪造虚假交易,给同伴正面评价,而给正常节点负面评价.因为在 P2P 网络,不存在仲裁机构,所以很难分辨交易的真假.通常,可以通过注册时严格的身份认证抑止 Sybil 攻击,但是信任系统仍然需要使用某种方法来验证节点之间的交易是否确实存在.

3 AARep 机制

本文设计了一种基于相似度的局部声誉机制 AARep,该机制提高了信任值计算的准确性和适应能力,并给出了一种轻量级的交易验证协议.在请求服务前,节点计算所有候选服务提供节点的信任值.为此,对每个候选者,请求节点首先搜索出候选者的客户作为推荐者;然后,通过置信因子和推荐节点与请求节点的相似度,加权推荐节点对候选者的本地评价,计算出信任值;之后,选择信任值最大的候选者进行交易;最后,客户端提交对本次交易的评价.

3.1 信任值定义

本节首先给出信任值和相关概念的定义,然后讨论 AARep 如何解决第 2 节所描述的各种攻击.

3.1.1 本地信任值

本地信任值 L_{ij} 直接来自于节点 i 对节点 j 的历史评价.假设节点 i 作为客户端,与节点 j 进行了 n_{ij} 次交易.定义节点 i 对节点 j 的第 k 次评价 r_{ij}^k 为

$$r_{ij}^k=\begin{cases}1, & \text{satisfactory}\\ 0, & \text{unsatisfactory}\end{cases} \tag{1}$$

其中,k=1,2,…,n_{ij}.

如第 2.1 节所述,聚集反馈节点可能在积累足够多的正面评价后去攻击其他用户.为了抵御这种攻击,信任

值应该及时反映节点的当前状态.本文使得本地信任值随着交易次序而衰减,减少较早交易的权重,增加最近交易的权重,即

$$L_{ij}=\sum_{k=1}^{n_{ij}}(\lambda^{n_{ij}-k}\cdot r_{ij}^{k})\Big/\sum_{k=1}^{n_{ij}}\lambda^{n_{ij}-k} \tag{2}$$

其中,$\lambda(0<\lambda\leqslant 1)$为衰减因子,反映衰减的强度.$\lambda$越小,本地信任评估就越能反映节点最近的行为.

3.1.2 相似度

相似度的计算基于两个节点对公共交易节点的反馈距离.令 CS_{ik} 表示与节点 i 和节点 k 都进行过交易的公共节点集合,节点 i 和节点 k 对节点 $l(l\in CS_{ik})$的反馈距离 D_{ik}^{l} 可以通过下式计算:

$$D_{ik}^{l}=|L_{il}-L_{kl}| \tag{3}$$

我们可以定义节点 i 和节点 k 的距离 D_{ik} 为 C_{ik} 上所有反馈距离的简单平均,但是进一步分析发现,每个 D_{ik}^{l} 都有不同的交易背景.通常,人们更原意相信经验丰富的节点所给出的反馈.由于两个节点与同一服务提供者之间有不同的交易次数,我们选择两者的平均值.于是,可以定义 D_{ik} 为

$$D_{ik}=\sum_{l\in C_{ik}}\left(D_{ik}^{l}\cdot\frac{n_{il}+n_{kl}}{2}\right)\Big/\sum_{l\in C_{ik}}\frac{n_{il}+n_{kl}}{2} \tag{4}$$

其中,n_{il} 和 n_{kl} 分别为节点 i 和节点 k 向节点 l 请求的交易次数.

节点 i 和节点 k 的相似度 S_{ik} 定义为

$$S_{ik}=1/D_{ik} \tag{5}$$

注意:节点 i 与自己的相似值为 1.

3.1.3 信任值

信任值表示节点关于其他节点能够正确地、非破坏性地完成交易的可能性的主观期望.如果所有节点都是可信的,信任值应该为成功交易次数与总交易次数的比值.令 T_{ij} 为节点 i 对节点 j 的信任值,其定义见公式(6).其中,节点 k 为与节点 j 交互过的节点,n_{kj} 为节点 k 向节点 j 请求过的交易次数.

$$T_{ij}=\sum_{k}(L_{kj}\cdot n_{kj})\Big/\sum_{k}n_{kj} \tag{6}$$

考虑到相似度,得到增强的信任值公式为

$$T_{ij}=\sum_{k}(L_{kj}\cdot n_{kj}\cdot S_{ik})\Big/\sum_{k}(n_{kj}\cdot S_{ik}) \tag{7}$$

理论上,当某个推荐节点的交易次数非常大时,信任值会趋于该节点的评价;如果所有的相似度都为 1,信任值将收敛至公式(6).公式(7)的极端情况与常识一致(见公式(8)和公式(9)).

$$\lim_{n_{kj}\to\infty}T_{ij}=L_{kj} \tag{8}$$

$$\lim_{\forall k,S_{ik}\to 1}T_{ij}=\sum_{k}(L_{kj}\cdot n_{kj})\Big/\sum_{k}n_{kj} \tag{9}$$

3.1.4 置信因子

然而,T_{ij} 受许多因素影响,使得它的值随参数呈非线性变化.因此,本文通过置信因子对公式(4)和公式(7)进行了修订:

$$D_{ik}=\sum_{l\in CS_{ik}}(D_{ik}^{l}\cdot PD_{ik}^{l})\Big/\sum_{l\in CS_{ik}}PD_{ik}^{l} \tag{10}$$

$$T_{ij}=\sum_{k,S_{ik}\geqslant\theta}(L_{kj}\cdot PL_{kj}\cdot S_{ik})\Big/\sum_{k}(PL_{kj}\cdot S_{ik}) \tag{11}$$

其中,分别用置信因子 PD_{ik}^{l} 和 PL_{kj} 替代了 $(n_{il}+n_{kl})/2$ 和 n_{kj}.与包含所有推荐节点的相似度算法不同,公式(11)过滤了相似度低的节点($S_{ik}<\theta$).显然,去掉那些明显可疑的推荐是合理的.

置信因子 PL_{kj} 定义为

$$PL_{kj}=n_{kj}^{\alpha} \tag{12}$$

其中,0≤α≤1.当α=0 时,赋予每个推荐者同等的权重;当α=1 时,意味着所有推荐者都是诚实的,信任值的计算基于所有提交的反馈.

置信因子 PD_{ik}^{l} 定义为

$$PD_{ik}^{l}=(n_{il}+n_{kl})^{\beta} \tag{13}$$

其中,0≤β≤1.当β=0 时,相似度考虑了公共交易集合中的所有节点;当β=1 时,相似度考虑了节点提交的所有反馈.

还可以从另外一个角度来理解置信因子.信任值和相似度是节点对其他节点的可信赖性和相似性的判断.通常,人们根据相对信心或背景知识来作判断.置信因子可以被看作是人们对观测值的信心,类似于数理统计的置信水平的概念.

3.1.5 进一步分析

本节将讨论 AARep 如何对抗第 2 节所述的前两种威胁.虚假交易攻击通过下一节描述的交易验证协议来抵御.

对于聚集反馈,我们使得本地信任值随着交易次序而衰变,减少较早交易的权重,增加最近交易的权重.加入交易衰减因子后,显式地区分了交易的重要程度,及时地反映了服务节点的当前状态,提高了信任评价的动态适应能力,减少了聚集反馈攻击的影响.

对于合谋欺骗,考虑到合谋欺骗攻击对团伙内的给予好评,对团伙外的给予差评,团伙内节点与团伙外节点的相似度会很低,AARep 使那些相似并且交易次数多的节点得到更大的权重,可以遏制合谋欺骗攻击的影响.

表 1 结合节点的经验和相似度对节点行为进行了分析.其中,斜体是采用传统方法(仅使得相似节点获得较高权重)时节点的行为,正常字体是 AARep 在不同情况下采取的行动.

Table 1 Node behavior analysis

表 1 节点行为分析

		Similar	Dissimilar
Normal node	Experienced	*More weight*	Not applicable
	Inexperienced	Less weight	Removed by θ
Malicious node	Experienced	?	Removed by θ
	Inexperienced	Less weight	Removed by θ

从表 1 可以看出,与本领域的其他研究相比,AARep 在消除不相似节点、降低无经验节点的重要性方面效果更好.对于问号标注的情况,即老练的恶意节点为大多数节点提交正常反馈,而向合谋欺骗节点提交虚假推荐,下面的定理可以证明不可信推荐者的夸大反馈并不会影响 AARep 机制的性能.

定理 1. 假设节点 k 是可信节点,则 $PL_{kj}\cdot S_{ik}=n_{kj}^{\alpha}$.

证明:根据公式(5)、公式(10)、公式(12)、公式(13),可得:

$$PL_{kj}\cdot S_{ik}=n_{kj}^{\alpha}\cdot\left(1-\frac{\sum_{l\in CS_{ik}}D_{ik}^{l}\cdot(n_{il}+n_{kl})^{\beta}}{\sum_{l\in CS_{ik}}(n_{il}+n_{kl})^{\beta}}\right).$$

假设节点 k 是合作的,即所有来自节点 k 的反馈都是可靠的,

$$\forall l\in CS_{ik},D_{ik}^{l}=0,$$

则

$$PL_{kj}\cdot S_{ik}=n_{kj}^{\alpha} \tag{14}$$

□

定理 2. 假设节点 k 与 CS_{ik} 中的节点 j 合谋,但对 CS_{ik} 中的其他节点提供真实反馈.令 $C=\sum_{l\in CS_{ik},l\neq j}(n_{il}+n_{kl})^{\beta}$,如果$\alpha$=$\beta$,则 $\lim_{n_{kj}\to\infty}PL_{kj}\cdot S_{ik}=C$;如果$\alpha$<$\beta$,则 $\lim_{n_{kj}\to\infty}PL_{kj}\cdot S_{ik}=0$.

证明:$PL_{kj}\cdot S_{ik}$可被表示为

$$PL_{kj}\cdot S_{ik}=n_{kj}^{\alpha}\cdot\left(1-\frac{\sum_{l\in CS_{ik}}D_{ik}^{l}\cdot(n_{il}+n_{kl})^{\beta}}{\sum_{l\in CS_{ik}}(n_{il}+n_{kl})^{\beta}}\right)=n_{kj}^{\alpha}\cdot\left(1-\frac{D_{ik}^{j}\cdot(n_{ij}+n_{kj})^{\beta}+\sum_{l\in CS_{ik},l\neq j}D_{ik}^{l}\cdot(n_{il}+n_{kl})^{\beta}}{(n_{ij}+n_{kj})^{\beta}+\sum_{l\in CS_{ik},l\neq j}(n_{il}+n_{kl})^{\beta}}\right).$$

假设除了节点 j,节点 k 对于 CS_{ik} 中其他节点的反馈都是真实的,即

$$\forall l\in CS_{ik}\cap l\neq j,D_{ik}^{l}=0,D_{ik}^{j}=1.$$

那么,

$$PL_{kj}\cdot S_{ik}=n_{kj}^{\alpha}\cdot\left(1-\frac{(n_{ij}+n_{kj})^{\beta}}{(n_{ij}+n_{kj})^{\beta}+\sum_{l\in CS_{ik},l\neq j}(n_{il}+n_{kl})^{\beta}}\right)=\frac{C\cdot n_{kj}^{\alpha}}{(n_{ij}+n_{kj})^{\beta}+C}\tag{15}$$

如果$\alpha=\beta$,则

$$\lim_{n_{kj}\to\infty}PL_{kj}\cdot S_{ik}=C.$$

如果$\alpha<\beta$,则

$$\lim_{n_{kj}\to\infty}PL_{kj}\cdot S_{ik}=0.$$ □

图 1 给出了 $PL_{kj}\cdot S_{ik}$ 随 n_{kj}(见公式(14)和公式(15))的变化情况.如图所示,对于可信节点(α=0.9 曲线),权重 $PL_{kj}\cdot S_{ik}$ 随 n_{kj}^{α} 呈线性增长.但是,对于欺骗节点,当$\alpha\leqslant\beta$时,PL_{kj} 与 S_{ik} 负相关,所以很难简单地通过一个量的增加来抵消另一个量的减少.同时,为了避免无限的交易次数,在信任值计算时可以只选用特定时间内的数据(例如 1 个月或者半年),丢弃掉过期的数据.而且,所有的交易必须通过下一节将要讨论的交易验证协议,这使得攻击者必须付出相当大的代价才能夸大交易数目.当然,我们也可以设定一个阈值过滤掉明显异常的交易次数.

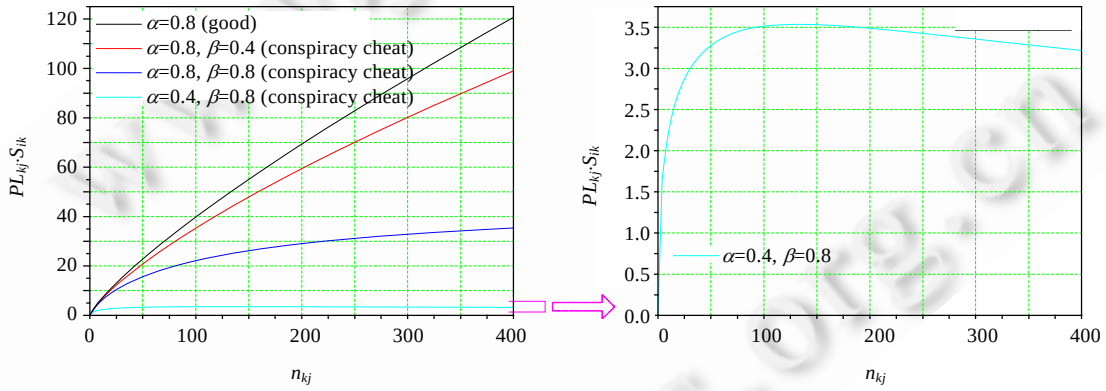


Fig.1 $PL_{kj}\cdot S_{ik}$ against n_{kj} (C=100)

图 1 $PL_{kj}\cdot S_{ik}$ 随 n_{kj} 的变化情况(C=100)

3.2 交易验证协议

针对虚假交易,为了使置信因子生效,提出了一种验证交易数目的交易验证协议.该协议实现简单,而且不需要任何第三方节点的参与.简而言之,接收者要求发送者随产品提供一张类似发票的证书,其中包含两者之间的交易次数.当需要验证时,节点出示自己的证书来证明交易的真实性.作为中间节点的档案点(见第 3.3 节)仅仅保存客户列表.

首先,假设在 P2P 文件共享网络中:

- 正常用户提供真实的文件和反馈;
- 恶意用户只企图分发恶意文件;

- 无论是正常用户还是恶意用户,作为请求者,节点都会提供反馈来影响其他节点的决定,并且不会篡改已有的评价.

假设节点 i 要计算节点 j 的信任值,为此,节点 i 需要计算和推荐节点 k 的相似度.于是,节点 i 要确认 CS_{ik} 中的节点 l 是否确实发送过文件给节点 k.假设节点 j 和节点 l 的档案点分别为节点 x 和节点 y.交易验证协议包括两个过程:(1) 伴随着文件下载的传输过程,发生在节点 k、节点 l 和节点 y 之间;(2) 发生在节点 i、节点 x 和节点 k 之间的验证过程.假设每个节点有一对公私钥.表 2 列出了本节用到的一些标记.

Table 2 Notation description

表 2 标记说明

Notation	Description
PK_i	Public key of node i
SK_i	Private key of node i
$MSG=(msg_1)\|\|(msg_2)\|\|\ldots\|\|(msg_n)$	Concatenate $msg_1,msg_2,\ldots,msg_n$ into message MSG
$[MSG]_K$	Use key K to sign the message MSG
$i--> j$	Node i send message to node j
$Cert_n$	The nth certificate

图 2 给出了传输过程.其中,$Feed_0$ 和 $Cert_0$ 设为 $NULL$.注意:如果节点 l 检查请求,发现证书无效,则节点 l 会终止本过程的执行.如果节点 k 没有收到签署的文件,或者签署的证书被破坏或缺失,则交易停止.

k:			$req\_msg=[(request)\|\|(Cert_{n-1})]_{SK_k}$
k	-->	l	req_msg
		l:	Check $cert_{n-1}$
		l:	$Cert_n=[(timestamp)\|\|(sender)\|\|(receiver)\|\|n]_{SK_l}$
		l:	$req\_msg=(file)\|\|(cert_n)$
k	<--	l	req_msg
k:			Check req_msg, if invalid, drop it
k:			If it is the first transaction with l, proceed as follows:
k:			$req\_msg=[(Cert_1)]_{SK_k}$
k	-->	y	req_msg
		y:	Check req_msg, if valid, update the customer list

Fig.2 Transmission process

图 2 传输过程

验证过程如图 3 所示.当节点 i 要求节点 k 出示与节点 l 交互的证据时,节点 k 发送节点 l 签署的最后一个证书给节点 i,其中包含了总的交易次数.节点 i 用节点 l 的公钥解密证书,检查证书内容.如果内容真实,则节点 k 为可信节点,否则不可信.

i:			$req\_msg=(req\_customer\_list\_for\_j)$
i	-->	x	req_msg
		x:	$req\_msg=(list)$
i	<--	x	req_msg
i:			Check the list and iterate every k in the list as follows:
i:			req_msg=(certification for the transaction with l)
i	-->	k	req_msg
		k:	$rep\_msg=[(L_{kl})\|\|(Cert_n)]_{SK_k}$
i	<--	k	req_msg
i:			Check the reply message

Fig.3 Validation process

图 3 验证过程

从传输过程和验证过程的描述可以看出,交易验证协议不需要任何第三方节点的参与.另外,交易验证协议

开销很低,是一种轻量级的协议.交易验证协议的开销主要来自查询所需信息的成本,与交互消息数目和每个消息的通信成本有关.由图 2 和图 3 可知,传输过程和验证过程仅需 3~4 个交互消息就可以完成,交互消息非常少.而每个消息的通信开销是由 P2P 网络的搜索机制决定的,如果搜索采用 P-Grid[15]的分布哈希表(DHT)机制,则查询复杂度为 $O(\log N)$,N 为系统规模.

设想 3 种违背交易验证协议的场景:

1) 节点 k 声称与节点 l 进行了 n 次交易,但实际并非如此.那么在验证过程中,节点 k 就无法出示节点 l 签署的证书给节点 i.为了攻击节点 l,节点 k 必须下载文件、存储所有交易证书,但是这个代价是非常昂贵的;
2) 如果节点 l 是一个别有用心的节点,但因为有其签名的证书,它也不能否认交易的存在;
3) 档案点 y 也不能篡改被节点 k 和节点 l 签名的客户列表.然而,节点 y 仍然有机会伪造证书或者选择性地丢弃某些列表,不过这可以通过备份方案和大多数投票方式来解决.

3.3 实现策略

节点信任值通过分布式方法求解.网络中任意节点同时具有两个角色:它既是用户节点,同时也是若干个用户节点的档案点.档案点通过 DHT 机制,例如 P-Grid 来实现.当然,也可以使用其他 DHT 方法来提高搜索、数据备份和失效恢复的效率,但这些并非本文的关注点.

一般来说,作为客户端,节点不会篡改它对其他节点的评价.但是作为服务提供者,它可能想隐藏自己的客户信息.AARep 简化了档案点的功能,档案点仅保存其用户节点的客户列表.与其他方法比较,该方法减少了档案点的关键数据,因此更新过程更少、更容易复制,也减少了安全隐患.AARep 的总体结构如图 4 所示:

- 客户列表(customer list)保存节点的交易关系,通过 DHT 机制响应其他节点的查询.
- 数据库(database)维护交易证书和本地信任值数据.
- 交易入口(transaction portal)负责交易验证协议的执行.
- 处理引擎(process engine)驱动其他部分完成处理,包括触发交易过程和信任值计算等.在请求文件 f 之前,节点 i 执行 *RequestService*(i,f)来选择提供节点,*RequestService*(i,f)的算法描述如图 5 所示.

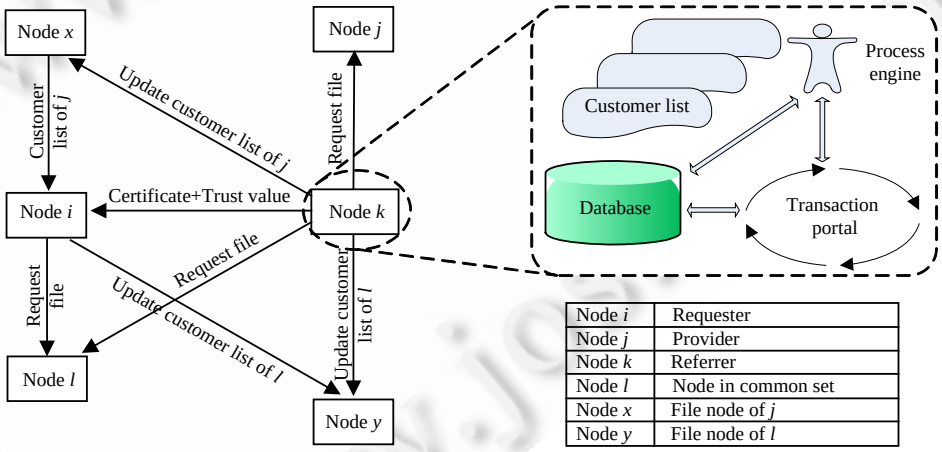


Fig.4 Framework

图 4 总体结构

```
Procedure RequestService(i,f )
i: requester
f: requested file
Begin
  Search f's owner
  for  each owner j  do
       Compute L_ij as per (1) and (2)
       Compute PL_ij as per (12)
       Compute j's file node x and request j's customer list from x
       for  each node k in j's customer list (except i)  do
            Get L_kj and verify relative certificate
            Compute PL_kj as per (12)
            for  each l∈CS_ik  do
                 Get L_kl and verify relative certificate
                 Compute  D_ik^l  as per (3)
                 Compute  PD_ik^l  as per (13)
            end for l
            Compute D_ik as per (10)
            Compute S_ik as per (5)
       end for k
       Compute T_ij as per (11)
  end for j
  Select the owner with the maximum trust value
End
```

Fig.5 Main process

图 5 主要处理流程

4 模拟实验结果

我们通过模拟实验来评价 AARep 对恶意攻击的抵御能力.作为参照,同时实现了平均值机制(MeanTrust,即信用值为所有节点对其评价的简单平均)和 PeerTrust(PSM/DTC Basic 和 PSM/DTC Adaptive)[10]机制.

4.1 实验环境

模拟实验环境为 PIV 2.66Ghz,1GB,实验代码用 C++语言编写.实验设想的应用场景为 P2P 文件共享应用,即用户的目标是下载其所需的文件资源.下载文件的真实性是其判断本次交易成功与否的唯一标准.

本文设计了两大类节点:

- 正常节点(S 类):这类节点无论在提供服务上还是在对其他节点的评价上都是真实的.
- 恶意节点(E 类):根据攻击方式不同,将恶意节点进一步分成 3 类:
  - ✓ 聚集反馈攻击节点(EA 类):这类节点在对其他节点的评价上是真实的,但在与其他节点交易时采取聚集反馈攻击行为.
  - ✓ 合谋欺骗攻击节点(EC 类):这类节点不仅提供不真实的服务,还联合起来形成恶意团伙,对团伙内的节点总是给出好的评价,对于团伙外的节点总是给出差的评价.
  - ✓ 虚假交易攻击节点(EF 类):这类节点通过虚假交易来吹捧同伴或诋毁正常节点.

在每个模拟周期内,每个节点从网络中其他节点处请求某个其不曾拥有的文件.为此,节点首先查找出拥有所需文件的所有节点,然后计算这些节点的信任值,选择信任值最大的节点进行交易;最后对服务节点进行评价,并在成功下载后把文件共享出去.

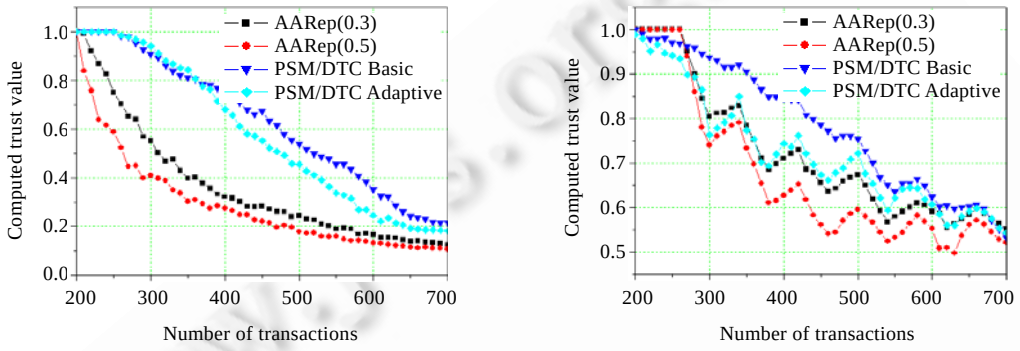
为了与 PeerTrust 对比(在 PeerTrust 的实验中,节点个数为 128),在本文的模拟实验中选取网络中有 128 个节点,2 000 个不同的文件.S 类节点以 95%的概率提供正常服务;在欺骗时,E 类节点以 5%的概率提供正常服务.每个节点初始化时拥有 100~150 个不同的文件,而且每个文件至少被一个 S 类节点保存.节点之间的初始信任值为 0.5.其他实验参数的配置如下:$\lambda=\alpha=\beta=0.8$,$\theta=0.3$(记为 AARep(0.3))或$\theta=0.5$(记为 AARep(0.5)).在 PeerTrust 的模拟过程中,基本时间窗口中的交易次数设为 200,自适应时间窗口中的交易次数设为 50.

4.2 实验结果

4.2.1 对抗聚集反馈攻击的能力

本节考察 AARep 机制对聚集反馈攻击的敏感程度,并与 MeanTrust 和 PeerTrust(PSM/DTC Basic 和 PSM/DTC Adaptive)机制进行比较.在下面的实验中,除了一个 EA 类恶意节点外,其他都为 S 类节点.参考 PeerTrust,我们模拟了两种情况:1) 恶意节点首先建立声誉,然后开始滥用它;2) 恶意节点在建立和滥用声誉之间振荡.实验中,节点随机选择其他节点进行交易,并选择一个 S 类节点定期地进行信任值的计算.

图 6 给出了节点 i 在不同场景下通过 AARep,MeanTrust 和 PeerTrust 方法计算得到的信任值.图 6(a)显示,当恶意节点通过一段时间的合作交互获得并积累一定的信任评价后,进行不合作时信任值的变化曲线.如图所示,这几种机制最终都导致信任值的降低.但是,AARep 对节点行为的改变更加敏感,当节点改变合作行为时,信任值迅速下降.而且,AARep(0.5)比 AARep(0.3)能够更快地检测出节点的恶意行为.图 6(b)显示,当恶意节点在建立和滥用声誉之间摇摆时,一个诚实节点对聚集反馈攻击节点的信任值变化曲线.同样,AARep 能够迅速发现节点的恶意行为,并延迟恢复信用值,特别是当θ=0.5 时.



(a) Node milking reputation
(a) 滥用声誉

(b) Node oscillating reputation
(b) 振荡声誉

Fig.6 Effectiveness against aggregate feedback

图 6 对抗聚集反馈攻击的效果

4.2.2 抵御合谋欺骗攻击的能力

下面的实验考察 AARep 机制抵御合谋欺骗攻击的能力,并比较了我们提出的机制与 MeanTrust 和 PeerTrust (PSM/DTC Basic 和 PSM/DTC Adaptive)机制的性能差异.在本节中,恶意节点为 EC 类.

图 7 为进行 700 次交易后,AARep,MeanTrust 和 PeerTrust 的失败下载率(fault download rate)随恶意节点比例的变化曲线.这里定义失败下载率为失败下载的次数占所有下载次数的比率,它直观地反映了信任机制的应用效果.从图 7 可以看出,随着恶意节点比例的增加,3 种机制的失败下载率不断增加.然而在所有情况下,AARep 的失败下载率都低于 MeanTrust 和 PeerTrust(而且增大θ带来更好的性能).这是由于 AARep 考虑了交易信息,过滤掉了可疑的节点,从而提高了信任值计算的准确性.值得注意的是,MeanTrust 机制在恶意节点较少时非常有效.这是因为在大多数情况下,通过相对排名,MeanTrust 就可以区分好、坏节点.但是当恶意节点达到 0.6 时,其失败下载率接近于 1.这表明,恶意节点能够完全愚弄 MeanTrust 机制,使其失效.

图 8 显示了当恶意节点比例一定时(0.4),AARep,MeanTrust 和 PeerTrust 的失败下载率随交易次数的变化情况.随着交互数目的增加,失败下载率逐渐下降.同样,与 MeanTrust 和 PeerTrust 机制相比,AARep 机制具有更高的准确性,因此具有更低的失败率.类似地,θ值较大时,失败下载率更低.

4.2.3 识别虚假交易攻击的能力

本节实验的目的是评价 AARep 机制识别虚假交易的能力.在下面的实验中,恶意节点属于 EF 类.

图 9 为经过 700 次交易后,AARep,MeanTrust 和 PeerTrust(PSM/DTC Basic 和 PSM/DTC Adaptive)的失败下

载率随恶意节点比例的变化曲线.当恶意节点增多时,这些机制的失败下载率都变大,其中,PeerTrust 的性能几乎呈线性下降.在所有情况下,AARep 的失败下载率都小于 MeanTrust 和 PeerTrust.而且,过滤掉更多低相似度的节点会产生更好的效果.AARep 的交易验证协议非常有效地识别了虚假交易.

当恶意节点比例一定时(0.4),图 10 显示了 AARep,MeanTrust 和 PeerTrust(PSM/DTC Basic 和 PSM/DTC Adaptive)的失败下载率随交易次数的变化情况.当交易数量增长时,失败下载率下降.同样,因为可以检测出虚假交易,AARep 机制具有最低的失败率.

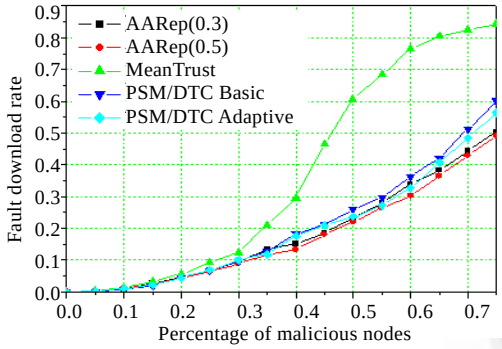


Fig.7 Impact of malicious node percentage on fault download rate

图 7 失败下载率随恶意节点比例的变化情况

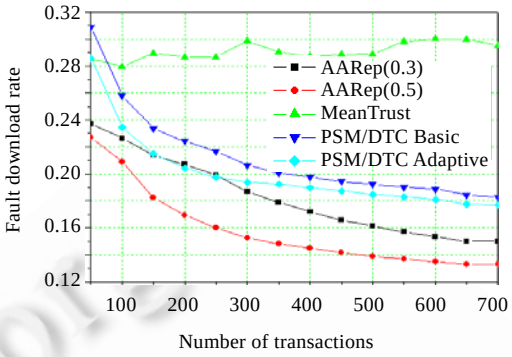


Fig.8 Impact of number of transactions on fault download rate

图 8 失败下载率随交易次数的变化情况

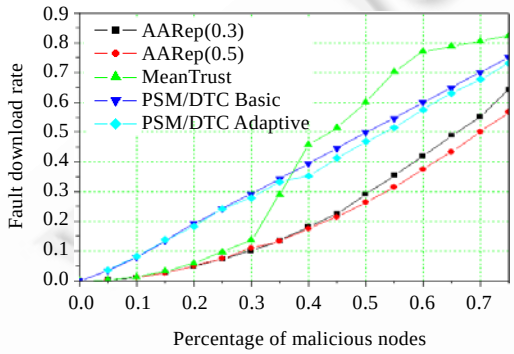


Fig.9 Impact of malicious node percentage on fault download rate

图 9 失败下载率随恶意节点比例的变化情况

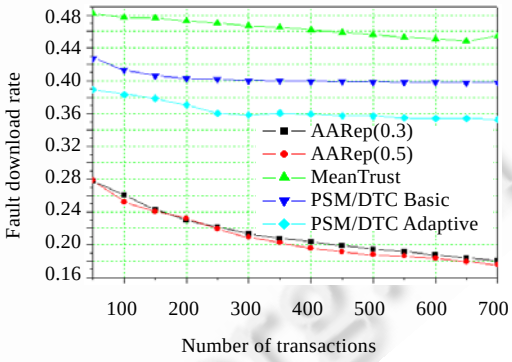


Fig.10 Impact of number of transactions on fault download rate

图 10 失败下载率随交易次数的变化情况

5 结 论

为解决现有 P2P 信任评价机制面临的聚集反馈、合谋欺骗和虚假交易等攻击问题,本文提出了 AARep 机制.与已有工作相比,该机制展现出以下两个显着特点:1) 信任值计算通过置信因子和推荐节点与请求节点的相似度来加权推荐节点的反馈,考虑了交易次数和衰减因子,提高了计算的准确性和适应能力;2) 为了验证交易信息,提出了一种简单的交易验证协议.仿真结果表明,AARep 机制对聚集反馈攻击较为敏感,能够抵御合谋欺骗攻击,有效识别虚假交易攻击.

References:

[1] KaZaA. http://www.kazaa.com/

[2] Liang J, Kumar R, Xi Y, Ross KW. Pollution in P2P file sharing systems. In: Makki K, Knightly E, eds. Proc. of the 24th Annual Joint Conf. of the IEEE Computer and Communications Societies (INFOCOM 2005). Miami: IEEE Press, 2005. 1174−1185. [doi: 10.1109/INFCOM.2005.1498344]

[3] VBS.GnutellaWorm. http://securityresponse.symantec.com/avcenter/venc/data/vbs.gnutella.html

[4] Buragohain C, Agrawal D, Suri S. A game theoretic framework for incentives in P2P systems. In: Shahmehri N, Graham RL, Carroni G, eds. Proc. of the 3rd Int'l Conf. on Peer-to-Peer Computing (P2P 2003). Los Alamitos: IEEE Press, 2003. 48−56. [doi: 10.1109/PTP.2003.1231503]

[5] Resnick P, Zeckhauser R, Friedman E, Kuwabara K. Reputation systems. Communications of the ACM, 2000,43(12):45−48. [doi: 10.1145/355112.355122]

[6] Jøsang A, Ismail R, Boyd C. A survey of trust and reputation systems for online service provision. Decision Support Systems, 2007, 43(2):618−644. [doi: 10.1016/j.dss.2005.05.019]

[7] Kamvar SD, Scholosser MT, Garcia-Molina H. The EigenTrust algorithm for reputation management in P2P networks. In: Lawrence S, ed. Proc. of the 12th Int'l Conf. on World Wide Web (WWW 2003). Budapest: ACM Press, 2003. 640−651. [doi: 10.1145/775152.775242]

[8] Dou W, Wang HM, Jia Y, Zou P. A recommendation-based peer-to-peer trust model. Journal of Software, 2004,15(4):571−583 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/15/571.htm

[9] Zhou RF, Hwang K. PowerTrust: A robust and scalable reputation system for trusted peer-to-peer computing. IEEE Trans. on Parallel and Distributed Systems, 2007,18(4):460−473. [doi: 10.1109/TPDS.2007.1021]

[10] Xiong L, Liu L. PeerTrust: Supporting reputation-based trust for peer-to-peer electronic communities. IEEE Trans. on Knowledge and Data Engineering, 2004,16(7):843−857. [doi: 10.1109/TKDE.2004.1318566]

[11] Zhai EN, Chen RC, Cai ZH, Zhang L, Lua EK, Sun HP, Qing SH, Tang LY, Chen Z. Sorcery: Could we make P2P content sharing systems robust to deceivers? In: Schulzrinne H, Aberer K, Datta A, eds. Proc. of the 9th IEEE Int'l Conf. on Peer-to-Peer Computing (P2P 2009). Seattle: IEEE Press, 2009. 11−20. [doi: 10.1109/P2P.2009.5284532]

[12] Zhang YC, Chen SS, Yang G. SFTrust: A double trust metric based trust model in unstructured P2P system. In: Yang Y, ed. Proc. of the 23rd IEEE Int'l Parallel & Distributed Processing Symp. (IPDPS 2009). Rome: IEEE Press, 2009. 1−7. [doi: 10.1109/IPDPS.2009.5161240]

[13] Srivatsa M, Xiong L, Liu L. TrustGuard: Countering vulnerabilities in reputation management for decentralized overlay networks. In: Ellis A, Hagino T, eds. Proc. of the 14th Int'l Conf. on World Wide Web (WWW 2005). Chiba: ACM Press, 2005. 422−431. [doi: 10.1145/1060745.1060808]

[14] Bahreman A, Tygar JD. Certified electronic mail. In: Proc. of the Internet Society Symp. on Network and Distributed Systems Security (NDSS'94). San Diego: Internet Society, 1994. 3−19.

[15] Aberer K. P-Grid: A self-organizing access structure for P2P information systems. In: Batini C, Giunchiglia F, Giorgini P, Mecella M, eds. Proc. of the 6th Int'l Conf. on Cooperative Information Systems. Berlin: Springer-Verlag, 2001. 179−194. [doi: 10.1007/3-540-44751-2_15]

附中文参考文献:

[8] 窦文,王怀民,贾焰,邹鹏.构造基于推荐的 Peer-to-Peer 环境下的 Trust 模型.软件学报,2004,15(4):571−583. http://www.jos.org.cn/1000-9825/15/571.htm

王淼(1975－),女,陕西乾县人,博士生,助理研究员,CCF 会员,主要研究领域为分布式计算.

张玉军(1976－),男,博士,副研究员,CCF 高级会员,主要研究领域为下一代网络.

陶飞(1987－),男,博士生,主要研究领域为分布式计算.

李国杰(1943－),男,博士,博士生导师,中国工程院院士,CCF 高级会员,主要研究领域为并行处理,计算机体系结构.