P2P 文件共享系统中基于访问控制的信任模型*)

左翠华　李瑞轩　卢正鼎
(华中科技大学计算机科学与技术学院　武汉 430074)

摘　要　在 P2P 文件共享系统中，节点之间共享彼此的文件。但是由于对等网络的匿名性和开放性，在文件共享系统中存在很多信任方面的问题，如恶意节点和自私节点的大量存在。为了孤立恶意节点和鼓励节点共享自己的文件，提出了基于访问控制的信任模型。首先，定义了一种信任机制，它引入了直接信任，推荐信任，时间衰减等因子。其次，对每个共享文件都加入了两个阈值——可信性总评价阈值和贡献量阈值来控制访问。对于每个请求者，只有当它有了对文件的访问资格后才能访问该文件。最后，通过大量的实验证明了本模型的可行性和高效性，特别是对恶意节点具有很好的孤立作用。

关键词　对等网络，访问控制，信任模型

Access Control Based Trust Model for P2P File Sharing Systems

ZUO Cui-hua　LI Rui-xuan　LU Zheng-ding
(School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, China)

Abstract　In peer-to-peer (P2P) file sharing systems, the participating peers share the files with others. However, due to anonymity and openness, there are many trust and free-riding problems such as malicious peers and free-riders in P2P file sharing systems. Proposed a remedy for these problems by introducing access control to trust model for P2P file sharing systems. First, the paper defined a general trust metric which integrates the following parameters: direct trust between the requestor and the provider, recommendation from other peers and time declining factor. Moreover, each shared file has two threshold values to control access. One is the credibility threshold value and the other is the contribution threshold value. For each requesting peer, it can not access a file until it satisfies accessing qualification of the file. Finally, a set of experiments prove the feasibility and benefit of access control based trust model. Especially, this model can isolate malicious peer effectively.

Keywords　Peer-to-peer, Access control, Trust model

1　引言

虽然 P2P 目前还没有被人们广泛接受的定义，但是它以其特有的特点成为计算机领域研究的热点。P2P 系统得到广泛应用和研究的一个很大的因素取决于 P2P 文件共享系统的成功，如 Napster, Gnutella 和 Kazaa。P2P 以其分布式的特性使得文件共享系统的鲁棒性、可用性和性能可随节点数目的增加而提高，系统的资源、带宽和能力也随着节点的加入而不断丰富和加强。此外，由于 P2P 文件共享系统中各节点之间的交互直接而对等，系统的资源可以高效地得到利用。目前，P2P 文件共享系统不仅仅局限于用户之间文件的传输，它的技术已经应用到企业和商业的平台[1]。尽管 P2P 文件共享的应用演变得越来越复杂，但是在 P2P 文件共享系统中对基于访问控制的信任模型的研究是相对比较少的。大多数的文件共享系统对节点是不加区分的，即所有的节点享受相同的下载文件权限，因此导致了系统中恶意节点和自私节点的大量存在。

本文提出了一种 P2P 文件共享系统中基于访问控制的信任模型。该模型的访问控制主要基于两个参数：可信性总评价阈值和贡献量阈值。可信性总评价阈值用来保证只有节点的可信性总评价值达到一定值才能从提供者上下载某个文件；贡献量阈值用来限定只有过去对自己做过一定贡献的节点才能从本节点上下载某个文件。模拟实验表明本模型简单可行，能有效孤立恶意节点，并且对节点间共享资源起到了很好的激励作用，有力地打击那些自私节点的行为。

本文的第 2 节介绍了相关工作；第 3 节介绍了基于访问控制的信任模型；第 4 节给出了访问控制策略；第 5 节展示了相关模拟实验结果；最后总结全文并指出了将来的工作。

2　相关工作

近来已经有很多文献提出了 P2P 环境下的信任模型，大致可以分为以下几类：

(1)基于 PKI 的信任模型；

(2)基于局部推荐的信任模型；

(3)基于数据签名的信任模型；

(4)基于角色的信任模型；

*)基金项目：国家自然科学基金项目(60403027, 60773191, 70771043)，国家高技术研究发展计划(863 计划)项目(2007AA01Z403)，中国博士后科学基金项目(20060400846)，湖北省自然科学基金项目(2005ABA258)，软件工程国家重点实验室开放基金项目(SKLSE05-07)，华为科技基金项目(YBIN2006089)。**左翠华**　博士研究生，研究方向为对等网络、分布式异构系统的安全；**李瑞轩**　博士，副教授，研究方向为分布式异构系统、分布式系统安全；**卢正鼎**　博士，教授，博士生导师，研究方向为分布式系统、智能信息系统、信息安全。

(5)基于 Bayesian 的信任模型;

(6)全局可信度模型。

基于 PKI 的信任模型引入了 CA 和 Leader Peer,因此这类系统往往具有中心依赖性并且难以扩展,eBay 就是这种模型的一个实例[2]。在基于局部推荐的信任模型中,节点一般采用简单的局部广播手段从有限的其他节点上获得某个节点的信誉度,这种信誉度往往是局部的和片面的。近几年里,很多对于 Gnutella 的研究就是基于这种模型的[3]。基于数据签名的信任模型根据下载该数据的用户对它的签名认证的多少来确定该数据的可信度。它只适合数据文件共享的应用环境,并且对集体欺诈行为没有防范能力,Sig2Dat 就是这种模型。在基于角色的信任模型中,节点依据其兴趣,加入到不同的社区,社区是拥有共同兴趣的节点集合。依据节点对于不同社区的隶属程度,决定其在不同方面的可信度,但该模型仍然存在一定的缺陷[4]。基于 Bayesian 的信任模型的核心思想是:依据文件质量、下载速度等参数,利用 Bayesian 概率的方法计算节点可信度,但这种计算实质上是基于用户自身的主观判定,往往具有片面性[5]。在全局可信度模型[6]中,为了获取全局的节点可信度,该模型通过邻居节点间相互满意度的迭代来获取节点的全局可信度,但是这种模型没有考虑到信任的不确定性,也没有考虑到惩罚因素和网络性能开销因素。Stanford 的 EigenRep 就是这种模型的一个典型例子[7]。此外,文献[8]提出了基于反馈的信任管理系统 PeerTrust,根据总体交易数,满意的交易数,以及平衡因子来刻画信任模型,通过节点间的协作来完成节点信任度的计算和存储,但是它没有区分过去交易和最新交易在计算信任度时的重要性。

上述各种 P2P 环境下的信任模型都存在自身的优势和不足。本文将提出一种基于访问控制的信任模型,并通过模拟实验来证明本模型的优势。

3 信任模型

这节主要讨论本模型的参数及对信任度的计算。

3.1 信任模型的参数

在本信任模型中,节点之间的信任度是取决于它们过去的行为。下面四个重要因素评估了节点的可信性:

(1)其它节点的推荐

两个节点在进行交易前,一般需要从邻近节点上获得对方的可信度信息来决定是否进行此次交易。对两个陌生节点而言,其它节点的推荐显得尤其重要。但对等网络中不可避免地会存在一些恶意节点,这些恶意节点可能对其它节点做出不诚实的评价,如恶意贬低或抬高某个节点的可信度。因此,在考虑其它节点的推荐信息时,推荐节点自身的可信性也是不可忽略的。

(2)直接信任因素

直接信任因素主要用于两个熟悉的节点间。它是熟悉节点间用来决定是否交易的一个重要因素,因为对比别人的推荐,节点总是更相信自己的经历。

(3)交易量因素

为了区分共享一个大容量资源和一个小容量资源所作贡献的不同,本文引入了交易量因素,即交易资源的大小。

(4)时间衰减因素

为了防止恶意节点周期性地欺骗行为,本文考虑时间衰减因素。当考虑到时间衰减因素时,一旦节点开始一次恶意行为,它的可信度就会大幅度的下降,这样就可以抑制节点进行恶意的欺骗行为。

3.2 信任计算

定义 1 定义 T_{ij} 作为节点 i 对节点 j 的直接可信性总评价。

$$T_{ij}=\begin{cases}\alpha t_{ij}+(1-\alpha)\dfrac{T_{ij}'}{1+\mu} & (0<t\leqslant T)\\ \beta t_{ij} & (t>T)\end{cases} \tag{1}$$

两节点交易完毕后,节点 i 会根据对本次交易的满意度给出一个可信性评价值 $t_{ij}\in[0,1]$。然后再根据式(1)来更新本地存储的直接可信性评价值 T_{ij}。T_{ij}'是过去节点 i 对节点 j 的直接可信性评价值。t 表示两个节点直接交易的时间间隔,如果节点 i 上没有存储和节点 j 的交易信息,则认为 $t>T$。T 是给定的一个周期值。如果两个节点相邻两次交易的时间差少于这个周期值 T,那么在计算新的直接可信性评价值时,以前的值是值得参考的。如果 $t>T$,那么它们过去的交易没有参考价值了,所以在计算新的直接可信性评价值时只考虑本次的评价 t_{ij}。此外,节点会周期性地删除那些没有参考价值的信息。α 和 μ 都是时间衰减因素,它们的值都介于 0 和 1 之间。α 表示在更新 T_{ij} 时本次交易与过去交易相比较的重要性,它的值越大则说明节点越重视最近的交易。μ 是过去交易记录的时间衰减因子。β 是一个折扣因子,避免陌生节点首次提供比较满意的服务就获得很高的直接可信性评价值。

定义 2 定义 R_{ij} 作为节点 i 根据与节点 i 和 j 都有交易记录信息的所有节点的推荐以及这些推荐节点的可信性所计算出的推荐可信性总评价值。当没有推荐节点时,则推荐总评价值为初始值 R_0。

$$R_{ij}=\begin{cases}\dfrac{\sum_{1}^{m}T_{ik}*T_{kj}}{\sum_{1}^{m}T_{ik}}\\ R_0\end{cases} \tag{2}$$

在式(2)中,k 为推荐节点,m 为所有推荐节点的总数。

定义 3 定义 C_{ij} 为节点 i 对节点 j 的可信性总评价,它由直接信任和推荐信任共同决定。当两节点间没有直接交易记录时,则仅考虑推荐信任。其中 λ 为权重因子,一般取值大于 0.5,因为对比其它节点的推荐,节点更相信自己的经验。

$$C_{ij}=\begin{cases}\lambda T_{ij}+(1-\lambda)R_{ij}\\ R_{ij}\end{cases} \tag{3}$$

定义 4 定义 D_{ji} 作为节点 j 从节点 i 上下载资源的总容量,亦称为节点 i 对节点 j 的贡献量。

$$D_{ji}=d_{ji}+D_{ji}' \tag{4}$$

其中,d_{ji} 表示一次交易中节点 j 从节点 i 上下载资源的容量。D_{ji}'表示过去节点 j 从节点 i 上下载资源的总容量。由于节点会周期性地删除本地一些过期信息,所以这里的总容量不可能包含以前所有的交易容量。

4 访问控制策略

在本模型中,节点 i 不需要保存每次交易的详细记录,而只需要在本地保存两个 XML 文件。一个是用来描述与节点 i 相识节点 j 的相关信息,其中包含节点 j 的 ID 和公钥信息、节点 i 和节点 j 最近的一次交易时间、节点 j 的直接可信性评价值 T_{ij}、节点 i 对节点 j 的贡献量 D_{ji}(即节点 j 从节点 i 上下载资源总容量,单位为 MB)。由于 D_{ji} 是存储在节点 i 上的,为了防止节点 i 对这个值进行非法修改,则 D_{ji} 是节点 j

用私钥加密后再发送给节点 i 的。节点 i 可以用节点 j 的公钥解密检查这个值是否正确，但不能进行修改。这个文件的例子如下：

```
〈TranInfo〉
  〈Transaction〉
      〈 PeerProperty  ID  =  '  ABC11311-D272-3EF3-84G4-5506555BBB66' PublicKey='XXXXXXXX' /〉
    〈TransactionDate Time='XXX' /〉
    〈Value Trust='0.38' Download='XXX' /〉
  〈/Transaction〉
  〈Transaction〉
      〈 PeerProperty  ID  =  '  983D3212-H742-3EFB-6CC3-4577752EE566' PublicKey='XXXXXXXX' /〉
    〈TransactionDate Time='XXX' /〉
    〈Value Trust='0.90' Download='XXX' /〉
  〈/Transaction〉
    ……
〈/TranInfo〉
```

另一个 XML 文件描述节点 i 自身属性及其共享文件的信息，它包含节点 i 的 ID，公/私钥、共享文件、贡献量阈值 D_{th}，可信性总评价阈值 C_{th}。其中，每个共享文件 F 都有对应的 D_{th} 和 C_{th}，这两个值的大小根据文件 F 的质量、级别来设置。当一个节点 j 需要下载节点 i 上文件 F 时，则必须满足下面两点：节点 j 对节点 i 的贡献 D_{ij} 不小于 D_{th}；节点 j 的可信性总评价值不小于 C_{th}。这个文件的例子如下：

```
〈PeerInfo〉
  〈PeerProperty ID='E4C43212-B367-6E48-78A3-4545452ABC34' PrivateKey='XXXXXXXX' PublicKey='XXXXXXXX' /〉
〈FileInfo〉
  〈SharingFile 〉XXXXXXXXX〈/SharingFile〉
    〈ContributeThreshold〉20〈/ContributeThreshold〉
    〈CredibilityThreshold〉0.80〈/CredibilityThreshold〉
  〈/FileInfo〉
  〈FileInfo〉
    〈SharingFile〉XXXXXXXXX〈/SharingFile〉
    〈ContributeThreshold〉0〈/ContributeThreshold〉
    〈CredibilityThreshold〉0.30〈/CredibilityThreshold〉
  〈/FileInfo〉
    ……
〈/PeerInfo〉
```

一般而言，一个刚加入网络的节点尽量把某些共享文件的阈值 D_{th} 和 C_{th} 都设置为最低值，以便其它节点能下载这些文件，从而提高自己的贡献量及可信性总评价，然后可以根据需要修改这两个阈值。对于某些高质量的文件，节点可以把 D_{th} 和 C_{th} 这两个阈值都设置得较高，只允许对自己有一定贡献量并且值得信任的节点来下载。节点为了能获得下载某些文件的权限，就需要共享本地的文件，从而提高自身的可信性总评价值和贡献量。可见，本模型对节点共享文件有很大的激励作用。

以节点 i 发出对文件 F 的请求为例，交易过程如下：节点 i 发起查询并获得一个提供者的结果集，这个结果集里包含了提供者的 ID 及下载文件 F 所需要的贡献量阈值 D_{th}；然后根据本地存储的对相识节点贡献量的信息来过滤掉一些提供者信息，这些提供者对文件 F 的贡献量阈值高于请求者 i 对它们的贡献量；再次，计算剩余提供者的可信性总评价值，发送下载请求给具有最高可信性总评价值的节点 j 并附上自身 ID、请求文件和对 j 的贡献量 D_{ji}，等待节点 j 的回复；接收来自节点 j 的回复，如果同意服务，则接收来自 j 的文件 F 并获得贡献量值 D_{ij}，根据对此次服务的满意度给出评价 t_{ij}，更新 T_{ij} 和 D_{ij}，并用私钥加密更新后的 D_{ij} 且发送到节点 j，如果是拒绝服务，则从结果集中删除节点 j，选择剩余提供者中具有最高可信性总评价值的节点，继续发送下载请求，按照上述步骤循环操作，直到下载到文件或者放弃下载该文件为止。

5 模拟实验

该模拟实验用来验证我们提出的信任模型的有效性。在模拟实验中，我们把节点分为两类：一类是诚实节点，它们提供真实可信的文件下载服务；一类是恶意节点，它们可能提供假文件下载服务，也可能在文件中附有病毒等恶意代码。我们假定一次交易的成功与否取决于请求者是否从提供者那里得到了真正想要的文件。实验参数如表 1 所示，我们将 1000 个文件随机且均匀地分配到 1000 个节点上，使得每个节点拥有 10 个不同的文件并且全部共享。每个文件的可信性总评价阈值和贡献量阈值都由共享该文件的节点来给定，因此不同节点上相同文件的访问条件是不一样的。

表 1 模拟实验参数

Parameter name	Parameter description	Parameter value
N	The number of peers	1000
M	The number of malicious peers	200
F	The number of files	1000
S	The number of files for each peer's owing	10
r	The rate of transactions a malicious peer acts maliciously	0.25
R_0	The initial value of total recommendation	0.3
T	The period time	2minute
q	The query frequency of each peer	5queries/min
α	Time declining factor	0.7
β	The discount factor	0.8
μ	Time declining factor	0.2
λ	The weight factor	0.7

5.1 交易成功率

本节将通过实验来显示本信任模型的交易成功率。在模拟实验中，每个节点都以相同频率 q 发送请求，当节点完成了 20 次下载时，该节点就停止发送请求，但仍然提供服务，直到所有节点都完成了 20 次下载后，实验结束。节点每次下载的目标是随机选择本地没有的文件。在这里，请求者从提供者上获得了文件就称为完成了一次下载，不管文件是否真实。下面我们按照 $r=0.25, 0.5, 1$ 来研究总交易数和交易成功率的关系，实验结果如表 2、表 3、表 4 所示（其它参数取表 1 中设置的值）。

表 2 $r=0.25$ 实验结果

Total transaction times	1000	5000	10000	15000	20000
Transaction failing times	50	232	447	528	540
Transaction success rate(%)	95	95.36	95.53	96.48	97.3

表 3 $r=0.5$ 实验结果

Total transaction times	1000	5000	10000	15000	20000
Transaction failing times	102	475	804	828	832
Transaction success rate(%)	89.8	90.5	91.96	94.48	95.84

表 4　$r=1$ 实验结果

Total transaction times	1000	5000	10000	15000	20000
Transaction failing times	199	904	1095	1132	1133
Transaction success rate(%)	80.1	81.92	89.05	92.45	94.34

由上述实验结果可知，当总交易数达到 1000，即平均每个节点完成一次下载时，交易失败的次数接近恶意节点数与其行骗比率 r 的乘积。随着总交易数的增加，交易失败次数的增加量越来越接近零。同时，随着总交易数的增加，交易成功率也不断的增加。表 4 显示了即使恶意节点以 $r=1$ 的比率行骗，当交易量达到 20000，即平均每个节点完成 20 次下载时，交易成功率就已经达到了将近 95%。

此外，恶意节点表现恶意行为的比率 r 越大，在起初阶段的交易成功率越低，但是随着总交易数的增加，交易成功率增加得越快，发现恶意节点的速度越快，交易失败次数增加得越慢。

可见，本实验不仅仅显示了本模型的交易成功率的增加幅度，从交易失败的次数可以看出本模型对恶意行为有很强地抑制作用。交易量越大，恶意节点行骗的机会越少。

5.2　孤立恶意节点的有效性

本节将通过实验来证明本信任模型对恶意节点的孤立作用。在本实验中，每个节点都以相同的频率 q 发送请求，当总交易数达到 20000 时，实验结束。假定恶意节点每次的行为都是恶意的，即 $r=1$，其它参数值为表 1 中给定的值。我们来分析一下恶意节点成功下载次数和总交易次数的变化关系。

表 5　实验结果

Total transaction times	1000	5000	10000	15000	20000
malicious peers' download times	206	974	1486	1760	1846
The rate of malicious peers' download times to total transaction times (%)	20.6	19.48	14.86	11.73	9.23

由表 5 可知，当总交易数为 1000 次，即每个节点平均完成一次下载时，恶意节点的下载次数占总交易数的 20.6%，与网络中恶意节点数占节点总数的比值 20%相近。当交易总量达到 20000 次时，恶意节点的下载次数占总交易数的比值不到 10%，并且在最后的 5000 次交易中，恶意节点的下载次数仅有 86。可见随着交易量的增加，恶意节点的身份逐渐被暴露，导致其下载文件越来越艰难，同时由上一节实验分析知总交易数越大，恶意节点越难行骗，所以恶意节点将会被快速地孤立起来。

结束语　在 P2P 网络中，文件共享系统得到了广泛的应用。在本文，针对文件共享系统中节点的欺骗和自私行为，我们提出了一种基于访问控制的信任模型。通过实验结果及分析，这种信任模型在下载成功率、对恶意节点的孤立以及激励文件共享等方面都有很大的优势。

参 考 文 献

[1] Park J S, Hwang J. Role-based Access Control for Collaborative Enterprise in P2P Computing Environment. SACMAT, Jun. 2003

[2] Resnick P , Zeckhauxer R . Trust among strangers in Internet transactions: Empirical analysis of eBay's reputation system // NBER Workshop on Empirical Studies of Electronic Commerce. California, 2000

[3] Cornelli F. Choosing reputable servants in a P2P network // Proceedings of the Int'l World Wide Web Conference. Hawaii, ACM Press, 2002: 441-449

[4] Khambatti M, Dasgupta P, Ryu K D. A role-based trust model for Peer-to-Peer communities and dynamic coalitions // Cole JL, Wolthusen SD, eds. Proc. of the 2nd IEEE Int'l Information Assurance Workshop. New York, IEEE Press, 2004: 141-154

[5] Wang Yao, Julita V. Bayesian network trust model in peer-to-peer networks // Moro G, ed. Proc. of the 2nd Int'l Workshop on Agents and Peer-to-Peer Computing. Berlin, Springer-Verlag, 2004: 23-34

[6] Zhang Zhen , Wang Xiao-Ming , Wang Yun-Xiao . A P2P global trust model based on recommendation // Proceedings of the 4th International Conference on Machine Learning and Cybernetics. Guangzhou, 2005: 3975-3980

[7] Kamvar S D, Schlosser M, Garcia-Molina H. Eigenrep: Reputation management in p2p networks // Lawrence S, ed. Proc. of the 12th Int'l World Wide Web Conf. Budapest. ACM press: 123-134

[8] Xiong Li , Liu Ling. PeerTrust : Supporting Reputation-Based Trust for Peer-to-Peer Electronic Communities // Proc. IEEE Trans. Knowledge and Data Engineering. 2004, 16(7): 843-857

(上接第 8 页)

[7] Liu Fang, Yao Li, Zhang Weiming, et al. A Conceptual Model of Agent Mediated Web Service[C] // 2004 IEEE International Conference on Services Computing, 2004. (SCC 2004). Proceedings. 2004, 9: 638-642

[8] Hui Wang , Yuhui Zhao , Deguo Yang , et al . An Agent-Based Services Composition Framework for Ubiquitous Media

[9] Vall'ee M, Ramparany F, Vercouter L. A Multi-Agent System for Dynamic Service Composition in Ambient Intelligence Environments. IEEE INTERNET COMPUTING, 2005: 11-12

[10] Huhns M N. Software Development with Objects, Agents, and Services

[11] Huhns M, Munindar, Singh P. A Semantic Web Services Architecture. IEEE INTERNET COMPUTING, 2005: 9-10

[12] Paul A, Buhler1, José M. Vidal2. Semantic Web Services as Agent Behaviors

[13] 杨欣，沈建京. 语义 Web 服务安全研究. 计算机科学，2007，34(2)

[14] 马长东，徐伟，李京. 一个基于 Agent 的 SOAP 应用框架. 计算机工程与应用，2004(1)：148

[15] 白伟华，苏卓夫，李吉桂. 服务代理在面向服务的体系结构中的应用. 计算机应用与软件，2006，23(11)

[16] Huhns M N. Agents asWeb Services. IEEE INTERNET COMPUTING, 2002: 7-8

[17] 戚玉松，钱柱中，是湘全. 基于 Agent 的 Web 服务组合研究. 南京理工大学学报，2006，30(3)

[18] Lu Hongen , Chhabra M . A Methodology for Agent Oriented Web ServiceEngineering

[19] Bozzo L , Mascardi V , Ancona D . COOWS : ADAPTIVE BDI AGENTS MEET SERVICE-ORIENTED COMPUTING

[20] Huhns M, Munindar, Singh P. Service-Oriented Computing: Key Concepts and Principles. IEEE INTERNET COMPUTING, 2005: 1-2