P2P 系统的信任研究*)

侯孟书 卢显良 周 旭 詹 川
(电子科技大学计算机科学与工程学院 成都610054)

摘 要 由于P2P系统的开放、匿名等特点,使得P2P系统对节点缺乏约束机制,节点间缺乏信任。针对以上问题,本文提出了一种新的P2P系统信任模型,该模型根据系统中节点的历史交易情况和系统中其它节点的推荐计算节点的信任度,节点根据计算的结果决定是否进行交易。仿真试验及分析表明,该模型能有效地评估节点的信任度,隔离恶意节点,提高下载成功率。

关键词 对等网络,信任,信任度,推荐

Study on Reputation in P2P System

HOU Meng-Shu LU Xian-Liang ZHOU Xu ZHANG Chuan
(College of Computer Science and Engineerinig, UESTC, Chengdu 610054)

Abstract Trust is a problem of particular importance in peer-to-peer system. However, the feature of the peer-to-peer system such as user anonymity, open nature makes that peers are not responsible for their irresponsible bartering history. This paper describes an algorithm to assess the peer which provides service based on the behavior of the peer. According bartering history and other peers recommend, the peer decides to download the sharing file from which peers. Analyses and simulations show that the model can assess the peer's reputation effectively, discard the malicious peer from peer-to-peer system and improve the rate of successful download greatly.

Keywords Peer-to-peer network, Trust, Reputation, Recommend

在P2P(peer-to-peer)系统中所有节点是对等的,这些节点既是客户机同时又是服务器,称之为对等机(SERVer-cliENT, SERVENT)。P2P系统一般分为三类:1)集中式的P2P系统,如早期的Napster;2)无中心的非结构化P2P系统,如Gnutella;3)无中心的结构化P2P系统,如Chord,CAN等。不论那类P2P系统,都是以节点间的协作为基础进行数据共享的,数据共享一般包括两个阶段:查找阶段和下载阶段。

节点间的信任是节点间协作的首要条件,然而由于P2P系统的开放、匿名、节点不为自身的行为担负责任等特点,导致P2P系统的服务质量(QoS)严重下降,很多查询结果的链接不可用,即使链接可用,也可能由于对方的随意离线,导致下载失败,更有甚者,恶意节点滥用P2P资源传播广告、病毒等文件来危害其它节点。针对以上问题,我们提出了一个新的信任模型,在该模型中,节点利用自己的交易经验和系统中其它节点的推荐,对和自己将要交易的节点进行信任度评估,根据评估的结果决定是否与被评估的节点进行交易,仿真试验表明,该模型能有效地将恶意节点从P2P系统中隔离出去,提高下载成功率。

1 信任系统

对于有中心的系统建立信任系统并非难事,像在线拍卖系统eBay的信任系统,在该信任系统中,买卖双方交易完成后,将对对方的评价提交给系统,系统通过记录交易成员最近6个月的交易情况,给出交易成员的信任度,系统依靠中心服务器存储和管理信任等级。然而在无中心的P2P系统中,由于没有系统中心服务器,对交易成员信任度的计算、存储和管理带来了困难。

文[1]是早期进行计算机信任模型建立的文献之一,该模型由于建立在信任的社会属性上,导致其过于复杂,很难实施。文[2]在文[1]的基础之上提出了在P2P系统中能够实现的方法,但是这种方法仍然需要维护一个复杂而庞大的数据结构,在实际情况下,维护这样一个数据结构是既费力又耗时的事情。文[3]的模型是在节点之间建立社会化的网络,并且参加者的信任度既来自其提供的服务又来自别人的推荐,每个节点拥有一个邻居节点的链表,并且通过和其它节点交换信息来更新节点的信任度。文[4]提出了一个二进制信任模型,节点仅仅存储与之交易过的不信任节点的信息,通过查询节点过去的行为获得节点的不信任度。文[5]针对P2P中的匿名带来的滥用P2P资源问题,提出了一种具有信任机制的P2P系统,节点在下载资源之前,通过分布式轮询其它节点来评价资源提供者的信任度。文[6,7]提出了类似的全局可信度模型,该类模型通过邻居节点间相互满意度的迭代,获取节点全局的可信度,节点与可信度高的节点交易。文[8]在分布式协作系统NICE中实现了一个分布式信任模型,当节点 i 和节点 j 首次交易时,根据信任节点的推荐进行交易,当交易完成后,节点 i 与节点 j 根据交易的情况相互给对方一个评价值,评价值以cookies的形式分别存储在节点 i 和节点 j 上,当节点 i 和节点 j 再次交易时,就可以根据原来的评价决定是否进行交易。文[9]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。文[10]提出了针对无中心非结构化的P2P系统的信任机制,通过节点的行为和节点的能力来评估节点的信任度。

*)电子信息产业发展基金资助项目,编号:[2002]11006.侯孟书 博士研究生,主要研究方向:分布式文件系统,P2P计算.卢显良 教授,博士生导师,主要研究方向:计算机网络,操作系统,信息安全.周 旭 博士研究生,主要研究方向:分布式文件系统,分布式存储.詹 川 博士研究生,主要研究方向:网络安全,邮件过滤.

2 信任度计算

2.1 信任度的相关定义

定义1 设 C_{ij}^t 表示在最近的时间段 t 内节点 i 对节点 j 的满意度，时间段 t 的引入，使满意度更能反映节点 j 的近期行为，该满意度刻画了节点 i 和节点 j 在最近的 t 时间段内交易的情况，设 $C_{ij}^t=\frac{S_{ij}^t}{S_{ij}^t+F_{ij}^t}$，其中 S_{ij}^t 为在最近的时间段 t 内，在节点 i 看来和节点 j 交易满意的次数，F_{ij}^t 为在最近的时间段 t 内，在节点 i 看来和节点 j 交易不满意的次数。例如当节点 i 从节点 j 下载文件 f 后，如果节点 i 满意则 $S_{ij}^t=S_{ij}^t+1$，否则，$F_{ij}^t=F_{ij}^t+1$。

定义2 设 T_{ij}^t 表示在最近的时间段 t 内，节点 i 对节点 j 的信任度，Φ_i 表示节点 i 的邻居的集合，n_{Φ_i} 表示节点 i 的邻居的个数。则：

$$T_{ij}^t=\begin{cases} C_{ij}^t & \text{在最近的时间段 } t \text{ 内，节点 } i \text{ 和节点 } j \text{ 有交易} \\ \sum_{k\in\Phi_j} T_{ik}^t T_{kj}^t/n_{\Phi_i} & \text{在最近的时间段 } t \text{ 内，节点 } i \text{ 和节点 } j \text{ 无交易} \end{cases}$$

显然，如果在最近的时间段 t 内，节点 i 和节点 j 没有交易，则 T_{ij}^t 的计算由一系列中间节点的推荐值相乘得到，这一系列中间节点，又称为推荐节点，将节点间的满意度作为推荐值，由节点 i、推荐节点和节点 j 组成了序列 $\langle i,k,\cdots,m,j\rangle$，序列的长度为计算迭代的深度，在本文的仿真试验中，设迭代深度为4。

定义3 设 M_i 为节点 i 的黑名单，用于记录在节点 i 看来的恶意节点。例如在节点 i 和节点 j 的交易中，如果节点 i 认为节点 j 对自己造成了危害（节点 j 提供了广告、病毒等文件），就将节点 j 加入 M_i。如果节点 $j\in M_i$，则将节点 j 的 ID（在 P2P 系统中的身份标识符）和节点 j 的 IP 地址加入 M_i 相应的数据结构中，在以后的交易中，节点 i 将不再和节点 j 进行任何交易。

2.2 信任度的计算

在 P2P 系统中，每个节点都与系统中的其它节点相连接，系统中所有节点及其链接形成一个覆盖在物理网络上的虚拟网络，该虚拟网络称为覆盖网（overlay network），在覆盖网中任意两个节点可以通信；消息通过网络中的边进行传播；节点之间地位平等。

当用户提交一个查询的时候，用户所在的节点首先检查本地是否有满足查询的结果，如果有查询结果，则向用户返回查询命中的消息；如果没有查询结果，则以某种策略沿覆盖网的边传播到邻居节点，当它的邻居节点收到查询后，进行同样的处理，以此类推。假设 P2P 系统的拓扑图如图1所示。

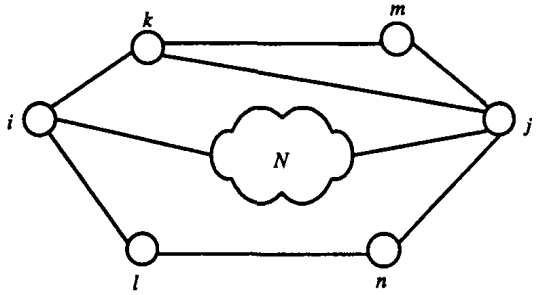

图1 P2P 系统拓扑图

图1中的子网络 N 表示从节点 i 到节点 j 的迭代次数大于设定的值4。当节点 i 收到查询结果后，如果节点 j 有满足查询的结果，这时节点 i 需要知道节点 j 的信任度，则对节点 j 的信任度计算如下：由定义2可知，当节点 i 和节点 j 在最近的时间段 t 内有交易时，则 $T_{ij}^t=C_{ij}^t$；当节点 i 和节点 j 在最近的时间段 t 内没有交易时，则 $T_{ij}^t=(T_{ik}^tT_{kj}^t+T_{il}^tT_{lj}^t+T_{iN}^tT_{Nj}^t)/3$，由于从节点 i 经过子网 N 到节点 j 的计算迭代次数大于4，故 $T_{iN}^tT_{Nj}^t=0$，所以 $T_{ij}^t=(C_{ik}^tC_{kj}^t+C_{il}^tT_{ln}^tT_{nj}^t)/3=(C_{ik}^tC_{kj}^t+C_{il}^tC_{ln}^tC_{nj}^t)/3$。

在计算 T_{ij}^t 时，当节点 i 和节点 j 在最近的时间段 t 内没有交易时，节点 i 向其邻居节点发出推荐请求，当推荐节点和节点 j 在最近的时间段 t 内有交易时，将满意度作为推荐值返回给请求节点，否则，推荐节点继续向它的邻居发出推荐请求，以此类推。当节点收到推荐请求后，将迭代次数减1，当迭代次数为0时，推荐请求就不再进行传播。

3 信任度算法

由于信任度的计算造成了网络流量的增加，可能影响 P2P 系统的性能，因此在信任度算法实现时，并不是每次查询后对所有满足查询条件的节点都进行信任度计算，在本文中，对节点的信任度设定生命期，只有当节点的信任度超过了设定的生命期时才进行信任度计算。信任度计算算法如图2所示。算法说明如下：

1）当节点 i 收到查询结果后，首先检查 j 是否属于 M_i，如果 $j\in M_i$，则不从节点 j 下载文件，否则，检查 T_{ij}^t 是否超出设定的生命期，如果没有超出生命期，则使用当前的 T_{ij}^t，否则，进行新的信任度 T_{ij}^t 的计算。

2）变量 *count* 用于记录本次计算的迭代次数，*count* 太小，得不到足够的推荐信息，*count* 太大，导致计算量迅速增大，影响 P2P 系统的性能。

3）*Neighbor*(*i*) 表示节点 i 的邻居节点的集合，*Num*(*Neighbor*(*i*)) 表示节点 i 的邻居节点的个数。

```
Reputation Algorithm
Input Peer i,Peer j,count
Output Tij
Float T(i,j,count)
{
  If--count<0Then Rreturn 0;
  If j∈Neighbor(i)then Return C[i,j];
  Else{
        For k∈Neighbor(i)
        Sum=Sum+C[i,k]*T(k,j,count);
        Return Sum/Num(Neighbor(i));
        }
}
```

图2 信任度计算算法

4 试验结果

为了评价本文提出的信任模型，建造了仿真试验环境，并且实现了没有引入信任模型的 P2P 系统。

网络模型：文[11]对 Gnutella 的测量和文[12]对 Freenet 的仿真都表明 P2P 系统中的节点间链接符合幂律（power-law）特性。在构建网络模型时，根据 P2P 系统的幂律特性设置节点的聚集度。

节点模型：将节点分成两类，一类是信任节点，信任节点在 P2P 系统中具有良好的行为，当其它节点从信任节点下载文件时，成功的几率较大。另一类是非信任节点，非信任节点又分两种，一种是搭便车（free loaders）节点，这类节点只从其它节点下载文件，而不共享自己的资源，在 Gnutella 中这种节点高达25%[13]，由于搭便车节点不共享任何内容，因此不会出现在返回的查询结果中，故在试验中没有仿真此类节点。

另一种是恶意(malicious)节点,这类节点要么提供虚假的文件,如广告、病毒程序等;要么经常终止下载,严重影响了P2P系统的性能。恶意节点最有可能进入黑名单,因此也最易从P2P系统中隔离出去。在仿真试验中,设定节点信任度的阈值ε,当$T_{ij}^t>\varepsilon$时,在节点i看来,在最近的时间段t内,节点j是可信任的节点。当$T_{ij}^t<\varepsilon$时,在节点i看来,在最近的时间段t内,节点j是非信任的节点。所有刚刚加入P2P系统的节点信任度为0。

在仿真过程中,节点可能处于在线,也可能不在线,根据文[14]对Gnutella的研究表明,50%的发起节点在线5小时以上,接近30%的发起节点在线24小时。

共享数据分布模型:P2P系统中共享数据的分布符合Zipf定律的结果,且Zipf指数的取值在[0.63,1.24]之间[15]。在仿真试验中,设定共享数据以文件的形式存在,且将文件分成50个种类,共计5000个,设定50个种类的复本数不同,受欢迎程度越高的文件,其在P2P系统中复本越多。

根据以上分析将仿真环境设置如下:

网络环境	节点总数	1000
	信任节点	50%～100%
	恶意节点	0%～50%
	节点的度	2～10
	节点查询传播的跳数	7
	节点的在线时间≥5小时	50%
共享文件	共享文件总数	5000
	共享文件种类	50
	共享文件分布	Zipf分布

根据以上的基本设置,首先对比了不同规模的恶意节点对基于信任的P2P系统(reputation system,没有引入黑名单机制)和没有引入信任模型的P2P系统(no-reputation system)的影响,结果如图3所示。可以看出,基于信任的P2P系统与没有引入信任模型的P2P系统相比,随着恶意节点的增加,引入信任模型的P2P系统下载成功率仍然维持在一个较高的水平,即使恶意节点达到50%,其下载成功率仍然接近60%。而没有引入信任模型的P2P系统在恶意节点达到50%时,其下载成功率仅仅接近30%。这是由于对于没有引入信任模型的P2P系统,其节点下载文件时,对所有的节点同等对待,而恶意节点要么提供的链接不可用,要么提供了虚假的文件,致使节点的下载成功率大大降低。

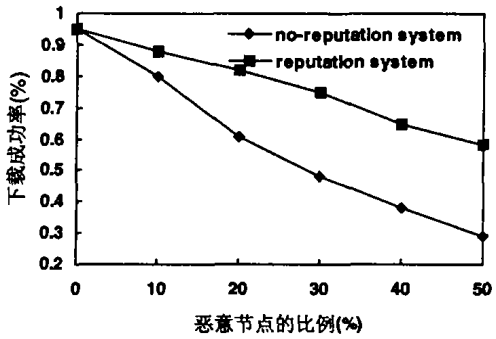

图3 不同规模的恶意节点对两种P2P系统的影响

其次,对比了不同规模的恶意节点对基于信任的P2P系统(reputation system)和引入黑名单机制的基于信任的P2P系统(reputation system with M)的影响,结果如图4所示。从图中可以看出,引入黑名单机制后,即使恶意节点的规模达到50%,带有黑名单机制的基于信任的P2P系统仍然具有较高的下载成功率,高达70%多,这是由于节点不再与列入黑名单的节点进行交易的结果,随着越来越多的节点将恶意节点列入黑名单,恶意节点就很难再有机会提高自己的信任度,这进一步加快了将恶意节点从P2P系统中隔离出去的步伐,虽然恶意节点存在于P2P系统中,实际上已经名存实亡,没有节点再与其交易。

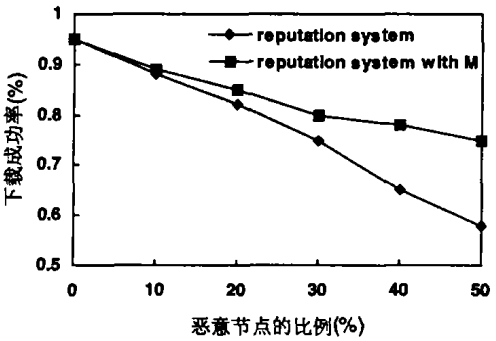

图4 不同规模的恶意节点对两种P2P系统的影响

结论 本文针对P2P系统中的信任问题提出了一个新的信任模型,在该信任模型中,节点根据自身的交易经验和其它节点的推荐对与之交易的节点进行有效的信任评估,以决定是否从评估的节点下载文件,通过分析和仿真说明,该模型能有效地评估节点的信任度,将恶意节点从系统中隔离出去,提高下载成功率。

参考文献

1 Marsh S. Formalising Trust as a Computational Concept:[Ph. D Thesis]. University of Stirling,1994

2 Abdul-Rahman A, Hailes S. Supporting Trust in Virtual Communities. In:Proc. of the 33rd Hawaii Intl. Conf. on System Sciences, 2000

3 Bin Yu, Munindar P. Singh: A Social Mechanism of Reputation Management in Electronic Communities. In:Proc. of the 4th Intl. Workshop on Cooperative Information Agents, Matthias Klusch, Larry Kerschberg (Eds.), Lecture Notes in Computer Science, Springer, 2000,1860:154～165

4 Aberer K, Despotovic Z. Managing trust in a peer-2-peer information system. In:Ninth Intl. Conf. on Information and Knowledge Management (CIKM), Nov. 2001

5 Damiani E, De Capitani di Vimercati, Paraboschi S, et al. A reputation-based approach for choosing reliable resources in peer-to-peer networks. In:9th ACM Conf. on Computer and Communications Security, Nov. 2002

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

7 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

8 Lee S, Sherwood R, Bhattacharjee B. Cooperative peer groups in NICE. In: IEEE Infocom, Apr. 2003

9 Xiong L, Liu L. Building trust in decentralized peer-to-peer communities. In:Intl. Conf. on Electronic Commerce Research (ICECR-5), Oct. 2002

10 Gupta M, Judge P, Ammar M. A reputation system for peer-to-peer Networks. In:NOSSDAV'03,2003. 144～152

11 Clip2 Company, Gnutella. http://www.clip2.com/gnutella.html

12 Hong T. In Peer-to-Peer: Harnessing the Benefits of a Disruptive Technology. edited by Andy Oram ～O'Reilley, Sebastopol, CA, 2001,14:203～241

13 Adr E, Huberman B A. Free riding on Gnutella:[Tech. Rep]. Xerox PARC 2000

14 Gnutella: To the Bandwidth Barrier and Beyond. http://lambda.cs.yale.edu/cs425/doc/gnutella.html. 2001

15 Saroiu S, Gummadi P K, Gribble S D. A measurement study of peer-to-peer file sharing systems. In:SPIE Conf. on Multimedia Computing and networking(MMCN), Jan. 2002