DOI:10.16707/j.cnki.fjpc.2017.11.067

一种基于改进推荐信任的 P2P 网络安全模型

李 珊[1]，吕学伟[2]
(1 淮安市经济和信息化委员会 江苏 淮安 223001；
2 淮阴师范学院计算机科学与技术学院 江苏 淮安 223300)

【摘 要】本文对基于推荐信任的全局模型进行了改进，为其引入了“兴趣群组”的概念。在该模型中节点和兴趣群组内的节点进行交易，提高了交易的安全性和服务质量；兴趣群组动态的更新组内成员，克服了局部交易的局限性和片面性。实验结果证明该模型能很快的识别恶意节点，有效的遏制节点的恶意行为，使 P2P 网络环境更加安全、可靠。

【关键词】P2P 网络；兴趣群组；推荐信任；安全模型

0 绪论

近年来 P2P(Peer-to-Peer)技术在文件共享领域内的应用，以及由此引发的社会问题使得 P2P 技术再一次被广泛关注，重新成为了 Internet 的研究热点。在 P2P 网络中节点之间直接进行信息和资源的交换，无需通过中间方，所有的节点其地位都是平等的，真正实现了网络间的平等沟通，有效降低了 C/S、B/S 类型网络中由于服务器产生的瓶颈效应，因此其在分布式计算、协同作业、对等计算等方面有着广泛的应用。虽然 P2P 技术有着诸多的优点，但与其缺点也是不可忽视的，其在安全方面的表现尤为突出，不可靠的服务质量、不诚实的交易行为、大量的欺诈行为，这些都给用户的安全造成了威胁。

由于 P2P 网络的自治性、匿名性、开放性等特点使得 P2P 网络既要面对传统的数据安全和网络安全，更要面对由于监管缺失和复杂结构造成更多的安全威胁，传统的管理方式和安全策略不适应于 P2P 网络，因此需要针对 P2P 网络结构的特点建立全新的安全机制。在 P2P 网络中，每个节点都是平等的，如果把人类社会中的每个人都看成一个节点，那么整个人类是会就是一个大的 P2P 网络，可以在借鉴人类社会的信任机制来建立 P2P 网络的信任安全模型，因此信任和信任机制被引入了 P2P 网络中。人们提出了各种 P2P 环境下的信任模型，例如：基于 PKI 的模型[1]、全局可信度模型[2]等等，每种模型都有优点，也有缺点，本文在全局信任模型中，引入了“兴趣群组”的概念来平衡全局可信度模型的全面性和局部推荐模型的高效性，以期获得更好的效果。

本文提出的改进推荐信任模型是吸收了全局信任模型和局部信任模型的优点，使节点能够在一个对特自己而言相当全局范围的节点所组成的兴趣群组中进行交易，这就兼顾了交易的高效性和全面性；此外模型对信任度的算法进行了改进，使得节点的交易更加安全、可靠。

1 模型的概念

定义 1 局部信任度[3]：一个节点根据以往的历史交易对另一个节点的可信程度进行的评价值，其定义如下：

$$P_{ij}=\frac{S_{ij}}{N_{ij}} \tag{1}$$

其中，P_{ij} 是节点 i 对节点 j 的局部信任度，S_{ij} 是在历史交易中节点 i 和节点 j 成功交易的次数，N_{ij} 节点 i 与节点 j 进行交易的总次数。

定义 2 推荐信任度[3]：一个节点根据历史交易记录给出的关于另一个节点的可信程度，其定义如下：

$$R_{ij}=\frac{S_{ij}-F_{ij}}{\sum_{k} S_{kj}} \tag{2}$$

其中，R_{ij} 表示节点 i 对节点 j 的推荐信任度，F_{ij} 表示在历史交易中，节点 i 和节点 j 交易失败的次数，如果 $\sum_{k} S_{kj}=0$ 则令 $R_{ij}=0$。

定义 3 内容相似度[4]：在整个网络中节点只和有限的节点进行交易，这些节点具有相同或者相似的内容，因此需要构造一种模型来计算节点间内容是否相近，具体如下：

假设整个 P2P 网络中的内容可以分为 m 种，分别用 C_1，C_2 ……，C_m 来表示，节点可以任意选择感兴趣的分类。分类向量 $V_i=[a_1,a_2,\cdots\cdots a_m]$ 表示节点 i 对网络中的内容感兴趣的情况，当节点对第 k 种分类感兴趣的时候，用 $a_k=1$ 表示，否则用 $a_k=0$ 表示，其中，$0<k<m$。

节点 i 和节点 j 的内容相似度可以用 $\vec{V}_i$ 与 $\vec{V}_j$ 夹角的余弦表示，其定义如下：

$$Simi_{ij}=\frac{\vec{V}_i \bullet \vec{V}_j}{\left\|\vec{V}_i\right\| \bullet \left\|\vec{V}_j\right\|}=\frac{\sum_{k=1}^{m} a_k \bullet b_k}{\sqrt{\sum_{k=1}^{m} a_k^2} \bullet \sqrt{\sum_{k=1}^{m} b_k^2}} \tag{3}$$

定义 4 兴趣群组：在人类社会中，具有相同的职业、爱好、兴趣等特征的人群间的交往更密切和频繁，因此在 P2P 网络中的节点应该也有相似的特性，因此我们对 P2P 网络中的节点的内容进行分类，根据不同类型的内容形成不同的兴趣群组，一个群组中的节点关心大体相同的内容，其交易也更频繁，通过节点间的内容相似度可以判断出节点是否在同一个群组中，如果节点的内容相似度大于某一个阈值的时候，我们就认为其具有相同的兴趣，可以组成兴趣群组，P2P 网络中的每个节点都可以归属于一个或者几个兴趣群组。本文中判断节点是否是同一群组的内容相似度阈值取值为 0.3。

定义 5 推荐可信度：一个节点提供的关于其他节点的推荐信任度的准确率，节点提供的推荐信任度和节点的真实比较接近，说明该节点提供的信息是可靠的，否则就认为其提供的信息不可靠。

文献 6 中用全局可信度作为推荐可信度，但是全局可信度仅仅代表了节点从事交易时在交易中的信誉，不能代表它就会诚实的对其它节点进行评价，因此我们需要引入一个新的衡量指标来代表节点做出的评价的可信程度。判断节点提供的推荐

评价是否准确，通过节点提供的推荐评价和其所在群组的其它节点提供的评价的平均值进行比较，以两者的差值来计算节点的推荐可信度。

用 C_x 表示节点 X 向节点 U 提供的关于节点 V 的推荐可信度。有如下计算公式：

$$c_x = c_{x'} + \frac{D_{ux} - \theta}{\lambda^{D_{ux}-\theta}} \quad (4)$$

$$D_{ux} = \frac{\sum_{j\in\Phi(u,x)} \left| R_{uj} - R_{xj} \right|}{N_{\Phi(u,x)}} \quad (5)$$

其中公式 (4) 中 C_x 为节点 U 原先的推荐可信度；ϑ 为节点所能接受的的最大偏移度，其中 D_{ux} 为节点 U 和兴趣群组内其它节点所做出评价的偏移量；λ 为增量控制因子，为大于 1 的数值。本文中 λ=2；

公式(5)中 $\Phi(u,x)$ 表示群组中与节点 U 和节点 x 都发生过交易的节点的集合。j 代表 $\Phi(u,x)$ 中的一个节点，其中，根据公式 2 可以计算出 R_{xj} 和 R_{uj} 的值，R_{xj} 为节点 V 提供的节点 j 的推荐信任度，R_{uj} 表示节点 U 提供的节点 j 的推荐信任度。$N_{\Phi(u,x)}$ 表示 $\Phi(u,x)$ 中节点的个数。

定义 6 不可信度：衡量节点进行恶意交易行为或者恶意提供虚假推荐可信度的指标。

在 P2P 网络中，一些节点通过大量的可信行为来获得好的信用，之后从事少量的恶意行为，然后接着再从事可信行为继续积累信任度，重复循环这样的行为，这类节点对 P2P 网络形成了潜在的安全隐患，但是仅用推荐可信度和信任评价是无法有效的制止节点的恶意行为，更不能辨别和剔除这些恶意节点，为此定义了不可信度这一指标，不可信度表示了节点进行恶意交易或者提供虚假推荐信息的程度，和节点进行的不可信行为的次数有关，是节点长期行为表现的一个衡量指标。其定义如下：$M_i=\beta k$，其中 M_i 是节点的不可信度，β 为惩罚因子，本文中 β=1。k 为节点的不可信行为的次数。不可信度和节点的不可信行为的次数直接相关，只要节点进行了不可信行为，其值就会增大，并且不会受节点其它行为的影响而减少，这就如同人类社会征信系统的“不良信用记录”一样，一旦有了就不会消失，会永远伴随节点而存在，因此不可信度可以有效的减少了节点的摇摆行为。

2 信任度的求解过程

信任度的求解分为以下几个步骤：

2.1 兴趣群组的初始化

当节点 i 加入 P2P 网络时，它不属于任何一个群组，因此其群组成员列表为空，它向网络中所有的节点发出查询消息 requestexit $(ID,\vec{V})$，向其它节点发送自己的内容向量 $\vec{V}_i$，其它节点在收到这个查询消息后，根据和自己的内容向量来计算和节点 i 的内容相似度，如果其值大于事先设定的阈值，则将其加入自己的兴趣群组列表，同时向节点 i 发出响应消息 responegroup (ID,Sim)，节点 i 收到响应消息后记录响应节点的 ID 和相似度，并比较相似度和阈值的大小来判断是否和相应节点属于同一群组. 当然恶意节点会传递虚假的内容相似度值给查询节点，以诱导查询节点，增加查询节点和其交易的机会，但是随着对节点的信任度的计算，恶意节点会被识别，通过群组的初始化，缩小了后续节点查询信任度查询的范围，提高了模型的响应速度。

为了判断一个节点是否可信，每个节点有一个可信标志，在初始状态下，假设所有的节点都是不可信的，因此其值都被设置为 0。由于内容相似度越高的节点其交易的概率越大，因此节点加入一个兴趣群组后，对群组的所有节点都按照相似度从高到底进行排序，优先和相似度高的节点进行交易，增加了其交易成功的概率，减少了信任计算的计算量，增加了交易的可靠性。

2.2 信任度的求解

在对一个节点进行评价获取其信任度的时候，需要结合自己查询节点和被评价节点的直接交易记录和兴趣群组中其它节点对被评价节点做出的评价值，这样才能对被评价节点有一个全面的准确判断，并且随着交易的进行，需要对这些值进行动态调整，这样信任度的值才能时刻反应一个节点当前的可信程度。因此可信度的计算公式如下：

$$G_{i,j} = \alpha \bullet \mathrm{P_{ij}} + (1-\alpha) \bullet \left(\frac{1}{N} \sum_{k=1}^{N} \mathrm{R_{kj}} \bullet \mathrm{C_k} \bullet \mathrm{simi_{kj}} \right) \quad (6)$$

其中 $\alpha=\frac{\gamma k}{k+1}$，$\gamma$ 为控制强度的因子，$0<\gamma<1$，k 为交易的次数。

2.3 兴趣群组的更新

由于在 P2P 网络中节点具有很大的随意性和自主性，其加入网络和离开网络毫无规律可言，P2P 网络时刻处于变化中，因此节点的兴趣群组也需要及时更新，否则无法保证群组中的节点仍然存在，容易造成以偏概全的现象，兴趣群组的存在和更新，吸收了全局信任模型的全面性，同时也兼顾了局部信任模型的高效性。

群组的更新分为以下几步：

(1)在进行交易时，请求节点会计算响应节点的综合信任度，在所有响应节点中寻找综合信任度值最大的节点进行交易，同时当综合信任度的值大于某一阈值的时候，就认为该节点是可信节点，修改其可信标志为 1，表明这些节点是可信的。

(2)请求节点向其兴趣群组中的所有节点发出简单查询信息，来确定群组中的节点仍在 P2P 网络中，如果节点在规定时间内发挥响应信息，表示该节点有效，否则表明该节点已经不在网络中，把其从群组成员列表中删除。

(3)随着交易的进行，节点之间相互了解的越来越多，节点所在兴趣群组的成员越来越多，为了提高信任度计算的速度，增强交易的可靠性，对群组的节点按照内容可信度和可信度进行重新排序。

(4)节点在离开 P2P 网络时需要向其所在兴趣群组中的其它节点发出的请求消息，请求退出，群组中其它节点在收到该消息后，从自身的群组成员列表中删除该节点，同时发出确认消息，之后请求节点离开网络。

兴趣群组的更新根据 P2P 网络中节点变化情况来决定，如果网络庞大，且网络中的节点经常变化，则就需要更高一点的更新频率，如果网络较小且节点变化少，则群组的更新频率可以适当降低一些。当然按照固定时间间隔或者定时更新也是一种更新策略。

在节点加入了兴趣群组后，节点几乎都和群组中的节点进行交易，该节点的交易成功与否极大取决于其兴趣群组中的节点，因此群组节点数据的正确性直接决定了其查询和交易的效率和效果，是模型的关键组成部分。

3 仿真及结果分析

为了验证模型的性能，我们进行了仿真实验，实验环境为：CPU I5-4590 3.3GHz，4G 内存，Windows 7 操作系统，仿真平台为李熊提供的 P2P 仿真器[5]，我们设计了多个实验来检测模型的性能和效果。设网络中共有 50 个交易节点，每个节点在发出 100 次下载后就停止继续交易。对每个节点，每次交易的时间间隔为 0.5 到2 之间一个随机数。网络中共发生 5000 次下载请求，系统共提供 1000 个文件，并将其平均分配到 5 个不同的内容分类中。为了检验本文提出的对模型进行改进后的效果，我们在上述仿真环境中实现了窦文提出基于推荐的 P2P 环境下的信任模型，对二者的性能进行了对比。此后窦文的模型被称为 RBPP-Trust 模型，本文提出的改进模型被称为 IGRBTrust 模型。

3.1 信任度计算的准确性分析

随机选取一个节点作为信任评价的对象，此外，随机选择 20 个节点，对其进行恶意评价。假设被评价节点 T 的真是信任度为 0.8，参数 γ=0.8，λ=2，综合信任度的阈值为 0.6，内容相似度的阈值为 0.3，不可信度的最大阈值为 20，恶意节点的评价位移阈值为 0.2，可信节点的评价位移阈值为 0.04。

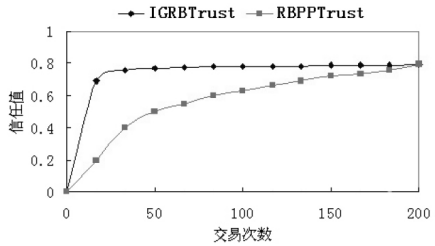

图 3.1 信任计算的准确性分析

由图 3.1 可见，在初始信任值为 0 的情况下，IGRBTrust 模型能迅速的逼近节点 T 的真是信任度，随着交易的进行，节点的信任度越来越接近真实值，对仿真数据进行分析后可以得出以下结论：

在 RBPPTrust 模型中节点综合信任度的计算包括了 P2P 网络中所有节点对被评价节点的推荐信任度，由于 P2P 网络中节点较多，且随时变化，因此其信任度需要较长的时间才能接近真实的信任度。在 IGRBTrust 模型中，由于推荐节点都来自于兴趣群组，其具有较高的内容相似度，因此其推荐信息具有较高的可信度，并且综合信任度的计算考虑了不可信度、动态比例因子等多个参数，使得综合信任度可以更快的接近被评价节点的真实信任度。

3.2 模型安全性分析

为了对模型的安全性进行检验，我们对网络中存在不同规模的恶意节点时，节点进行文件下载的情况进行了实验

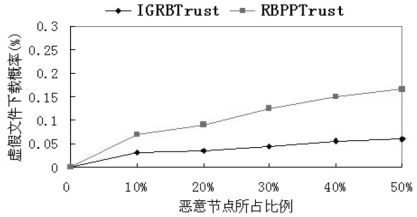

图 3.2 诋毁夸大情况下虚假文件的下载比率

实验结果如图 3.2 所示，由于两种模型都是基于推荐信任的模型，都征询了网络中其他节点的意见，从而保证了综合信任度的计算收到恶意节点的额恶意行为的影响较小，有效的保证了交易的安全性，不同之处在于，RBPPTrust 模型是通过对两次交易评价进行比较和交易双方都对彼此进行评价来控制节点的恶意行为，这种方法对节点的恶意摇摆行为是无效的，因此虚假文件的下载比例随着恶意节点的增加而增加。IGRBTrust 模型通过计算评价偏移度来计算节点的推荐可信度，这有效的拒绝了节点的不诚实评价行为，并且不可信度的引入在一定程度上控制了摇摆型节点的不诚实行为，因此在安全性方面，I-GRBTrust 模型的安全性有强于 RBPPTrust 模型。

3.3 请求信任度的消息转发分析

由于 RBPPTrust 是全局信任模型，因此其在计算综合信任度时需要在整个 P2P 网络范围内进行消息的迭代转发，严重的情况下将产生洪泛效应；而在 IGRBTrust 模型中，消息的请求和应答多发生在趣群组内，因此其消息转发的数量是比较小的。

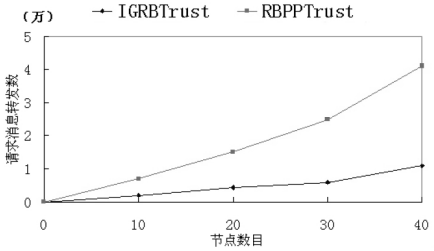

图 3.3 相同节点数目情况下请求消息的转发数

实验结果如图 3.3 所示，随着节点增多，两种信任模型转发的消息数量都有所增加，但是在 RBPPTRust 模型中，消息的增加是指数级的增加，这就决定了其在大规模网络中很难实现；在 IGRBTrust 模型中，即使节点大量增加，消息的数量基本保持在很小的范围内变化，因此 IGRBTrust 模型具有开销小，可扩展的优点。

结论

本文对窦文等人提出的基于推荐的信任模型进行了改进，新模型吸收了全局模型和局部模型的优点，兼顾了有效性和全面性，提高了安全性，具有工程可行性。实验结果证明这种模型取得了比较好的效果。

参考文献：

[1] 1Kamvar SD, Schlosser MT. EigenRep: Reputation management in P2P networks.In Proceedings of the 12th Int'l World Wide Web Conference,Budapest: ACM ress:123-134.

[2] 2 Vu Le-Hung, HauswirthM ,Aberer K. Towards P2P-based Semantic Web Service Discovery with QoS Support. In: Proceeding of Workshop on Business Processes and Services(BPS) , Nancy , France ,2005.

[3] 3 窦文，王怀民，贾焰等.构造基于推荐 Peer to Peer 环境下的 Trust 模型.软件学报.2004,15(4):571 一 583.

[4] 4 陈颖熙，李贤有，顾明等.基于内容相似度的对等网络信用模型研究. 计算机科学 2007,34 (8) 92-94.

[5] 5Li Xiong, Ling Liu. A Reputation-Based Trust Model for Peer-to-Peer Ecommerce Communities.In Proceedings of IEEE Conference of E-Commerce, ACM Press, 2003: 275-284.

作者简介：

李珊 (1980-) ，女，陕西宝鸡人，助理工程师，本科，研究方向为：信息安全；吕学伟 (1979-) ，男，陕西西安人，硕士，讲师，研究方向：信息安全。