

引入不确定性理论的 P2P 信誉模型研究

席菁¹, 胡平², 陆建德²

(1. 苏州大学计算机学院, 苏州 215006; 2. 江苏省计算机信息处理技术重点实验室, 苏州 215006)

摘要: P2P 系统匿名、动态的特点使得它成为各种自私和恶意行为的温床。为了解决因 P2P 的高度动态特性而引起的节点欺骗问题, 基于概率模型和模糊数学的基本思想, 构建一个 P2P 安全信任模型——BGTR, 实验证明, 通过建立 BGTR 模型, P2P 网络可以获得更高的安全性, 能够较好地解决冒名、协同作弊以及“搭车行为”等问题

关键词: 不确定性; 贝叶斯网络; 全局信任值

Research on P2P Reputation Model with Uncertainty Theory

XI Jing¹, HU Ping², LU Jian-de²

(1. School of Computer, Soochow University, Suzhou 215006; 2. Jiangsu Province Computer IT Key Lab, Suzhou 215006)

【Abstract】The anonymity and dynamic characters of P2P (Peer to Peer) system makes P2P become hotbed with selfish and vicious action. In order to solve the problem of node cheating caused by high dynamic property of P2P, this paper designs a security trust model——BGTR based on the idea of probability model and fuzzy mathematics. The P2P network with BGTR trust model proves higher security and it has better performance to resolve the problems of impostors, malicious collective attacks and free-rides.

【Key words】 uncertainty; Bayesian network; global trust value

1 概述

目前, 已经存在多个基于 P2P 环境的理论模型, 如文献[1]中提出的 EigenTrust 模型, 它提出了基于 DHT 的安全分布式计算方式, 要求节点在与其他节点进行交易时, 先查询本地表, 若查询不成功, 则向朋友节点查询、直至向朋友的朋友查询, 这种模型存在用户合作作弊提高自己信任值的问题, 并且该算法很难描述每个节点对系统整体的贡献值。文献[2]提出了 TPP 模型, 该方案中每个节点通过计算被查询节点信任值的方式, 使用信任向量建立本地信任表, 并提交对另外节点的评价以建立全局可信表, 最终建立一个信任网络。该模型能够较好地提高交易成功率, 但算法复杂度较其他模型也有所提高。文献[3]中提出的 Trustme 模型, 节点的全局可信记录存放于其他节点, 交易时先到存放节点上去查询, 它通过全局的迭代计算来获得这些节点的全局可信度, 避免了合作作弊现象, 但通信开销比较大。

本文在分析了这些现有的信誉模型的基础上, 提出了一种由主动发起查询的节点去计算被查询节点的可信值和被查询节点主动计算节点的全局可信值两者相结合的新型信誉模型 BGTR。

2 BGTR 模型的基本原理及其关键技术

2.1 P2P 动态信任机制的基本需求

P2P 网络具有高度的动态性、自治性和异构性, 这些特性使得为了更准确反映节点的真实情况, P2P 信誉模型需要满足如下基本需求:

(1)信任机制应对近期的交易经验给予更大的关注, 并且对于成功和不成功的交易, 信任值的变化应该给予不同的处理。

(2)信任机制在信任值计算上不能只专注于交易数目。

(3)计算信任值时应专注于近期的交易, 但也不能忽略

早期交易, 两者要综合考虑。

(4)信任机制要能够抵御有计划的欺骗行为, 如通过一段时间的成功交易获取较高的信任值后再进行恶意行为、合同作弊等。

2.2 BGTR 模型的基本框架及其关键技术

BGTR 模型采用全局信任值、直接交互值和推荐信任值三者联合考虑的方法, 通过判定树的逐级匹配, 构建出一个能够及时反映 P2P 网络变化的信誉模型, 基本框架见图 1。

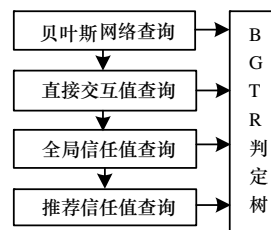


图 1 BGTR 模型框架

节点通过搜寻获得被查询节点后, 执行以下步骤:

(1)通过 Bayesian 网络进行查询, 判定被查询节点是否满足查询节点对资源提供者的一些基本要求, 若满足, 则转步骤(2), 若不满足, 表明该节点没有提供交易的能力, 重新选择下一个节点。

(2)根据自己与被查询节点以前的交易记录计算出直接交互值, 通过 BGTR 判定树进行判定: 若信任值达到判定树规定的阈值 T_1 , 则两者之间进行交易; 若信任值介于阈值 T_1

基金项目: 江苏省自然科学基金资助项目(BK2004039)

作者简介: 席菁(1984—), 女, 硕士研究生, 主研方向: 计算机网络, 信息安全; 胡平, 硕士研究生; 陆建德, 教授

收稿日期: 2008-09-10 **E-mail:** jingxi@ustc.edu.cn

和 T_2 之间, 则转向步骤(3); 若信任值小于 T_3 , 则重新选择被查询节点后, 转向步骤(1)。

(3) 查询由被查询节点主动计算的该节点的全局信任值并将它与步骤(2)中得到的直接交互值复合, 根据 BGTR 判定树进行判断: 若信任值大于阈值 T_3 , 则两者之间进行交易; 若信任值介于阈值 T_3 和 T_4 之间, 转步骤(4); 若信任值小于 T_4 , 则重新选择被查询节点后, 转步骤(1)。

(4) 利用基于权重及信任向量的推荐信任模型进行推荐信任值的叠加, 并将它与步骤(3)中得到的混合信任值综合考虑, 通过 BGTR 判定树的判定: 若信任值大于阈值 T_5 , 则两者之间进行交易; 若信任值小于 T_5 , 则重新选择被查询节点后, 转向步骤(1)。

BGTR 信誉模型为了满足 2.1 节中提出的要求, 在模型中加入了 Bayesian 网络模型判定、信任值的三元组表示、风险指数等一系列技术, 具体内容将在第 3 节中具体阐述。

3 BGTR 模型中的不确定性应用及其关键技术

3.1 贝叶斯网络判别模型

BGTR 模型为了提高系统效率, 减少通信开销, 引入了类贝叶斯模型, 如图 2 所示。P2P 节点提供的资源具有多样性, 每个节点的能力也是不一致的, 以文件共享网络为例, 有些文件提供者可以提供高速网络来传输文件, 而有些节点却只能提供低速网络; 有些文件提供者可能偏向于提供音乐文件, 并在提供这些服务时具有良好的信誉, 而另一些文件提供者可能偏向于其他类型的文件。类似于这样节点行为的一系列动态因素, 使得对于节点信任值不能以一种绝对的眼光看待, 应该根据不同的场合有不同的理解。现在的 P2P 网络中, 一些恶意节点利用节点的这种动态性, 通过在某一方面的行为骗取较高的信任值, 而利用已获得的信任度在另一方面进行欺骗行为。为了解决这类恶意行为, 在模型判定树的最顶端需要设置类 Bayesian 判别模型。

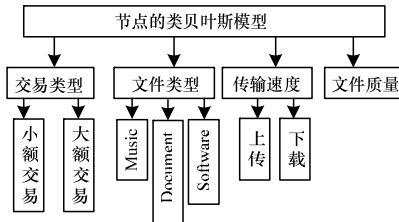


图 2 类贝叶斯网络结构

本文采用的类 Bayesian 模型如图 2 所示, 它根据节点自动搜寻到的以往交易节点对它的反馈值经过统计计算后所得的结果构建而得, 每一个节点都有一个属于自己的类 Bayesian 网络, 当一个节点希望从资源提供者处高速下载一个 document 类型的文件时, 它可以先查询资源提供者的类贝叶斯网络、获知它能否有提供这种要求的能力。类 Bayesian 网络所采用的统计原则如下所示:

$$p(h|e) = \frac{p(e|h)p(h)}{p(e)} \quad (1)$$

根据式(1), 以 document 类型的文件为例:

$$p(\text{document}) = p(\text{document} | T=1) + p(\text{document} | T=0) \quad (2)$$

其中, $p(\text{document} | T=1)$ 表示交易成功时类型为 document 的交易的失败概率; $p(\text{document} | T=0)$ 表示交易失败时类型为 document 的交易的失败概率。

$$p(\text{document} | T=1) = \frac{p(\text{document}, T=1)}{p(T=1)} \quad (3)$$

$$p(\text{document}, T=1) = \frac{m_{\text{document}}}{n} \quad (4)$$

$$p(T=1) = \frac{m}{n} \quad (5)$$

其中, m_{document} , m , n 分别表示交易次数中类型为 document 的交易的总的满意次数、交易过程中总的满意次数、总的交易次数。 $p(\text{document} | T=0)$ 的计算类似于 $p(\text{document} | T=1)$, 且满足:

$$p(T=0) + p(T=1) = 1 \quad (6)$$

3.2 信任值的三元组表示与风险指数的关系

在现有的主流信誉模型中, 对节点之间的交易往往只给予成功、不成功 2 种结论, 而在现实世界里, 一些事件往往不能给予非常肯定的结论。

本文利用(信任值, 不信任值, 不确定值)三元组来表示节点的信任值, 将现实世界的不确定性因素引入模型中, 利用激励机制, 对满意交易、不满意交易、不完全满意交易给予不同的奖惩策略; 并且将交易的风险指数与信任值中的不确定值联系起来, 风险指数越高, 不确定值在信任值计算中作为可信值所占的比例也就越大(表 1)。本模型中的不确定值与现实世界类似, 随着两节点之间的交互次数的增加, 两节点之间的了解也越多, 不确定因素越少, 不确定值也越小。

表 1 风险指数与不确定值的关系

风险指数 I	参与计算的可信信任值
0	信任值
1	信任值 + 0.2 × 不确定值
2	信任值 + 0.4 × 不确定值
3	信任值 + 0.6 × 不确定值
4	信任值 + 0.8 × 不确定值
5	信任值 + 不确定值

3.3 交易次数及交易时段对信任值计算的影响

现有的大部分信誉模型对节点信任值的计算都不能很好地适应节点快速转变的行为, 例如在文献[4]所提及的模型中, 信任值只能增加不能减少; 而在其他的现有典型模型中, 模型大部分都对节点的容忍性太强, 不能够及时反映节点的最新状态。本文所设计的 BGTR 模型将信任值的计算采用归一化处理, 即所有取值在 [0, 1] 区间内, 信任值的计算既考虑交易总量, 又给予近期交易更大的权值, 如式(7)所示:

$$T(a) = \Phi(\alpha \times E(_, a), \beta \times E(x, a) | x \in \text{与节点 } a \text{ 最近交互的节点}) \quad (7)$$

其中, $E(_, a)$ 表示节点 a 进行的所有交易, 在本文中直接交易值以及全局信任值均采用式(7)进行计算, 实现见 3.3.1 节和 3.3.2 节。

3.3.1 直接交易值的计算

节点的直接交易值, 可以采用本地信任表中存储的数据来进行计算。节点之间进行交互, 会将一些表现比较友好的节点的交易记录保存, 作为自己的朋友节点。BGTR 模型采用一种定长的二进制向量来记录这些交易情况, 本地信任表中一共使用了 2 个向量, 向量 1 用来记录成功和失败的交易的情况, 向量 2 用来记录交易中带有不确定性因素的交易情况。假设 2 个向量均采用 8 bit 的二进制来表示: 节点 A 与节点 B 进行完一次交易, 若交易结果明确, 则将向量 1 左移一位, 如果交易成功, 则将右边空出的一位填 1, 如果不成功, 则填 0, 同时向量 2 左移一位, 右边空出的一位填 1; 如果交

易情况介于失败与成功之间，则将向量 1 与向量 2 均左移一位，右边空出的一位填 0。假设节点 A 对节点 B 的最近交易情况为 11010100，交易状态为 11011101，交易完成后可能出现 3 种情况，如图 3 所示。

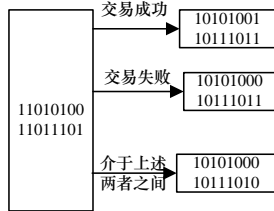


图 3 交易状态变化

对于三元组中数值的计算，以交易前为例，计算方法如下：

$$\text{信任值} = \frac{\text{向量1中的有效个数}}{\text{向量1的有效位数}} = \frac{4}{8}$$

$$\text{不信任值} = \frac{(\text{向量1} \cup \text{向量2}) \text{ 中的有效个数} - \text{向量1中的有效个数}}{\text{向量1的有效位数}} = \frac{(\text{11010100} \cup \text{11011101}) \text{ 中的有效个数} - (\text{11010100}) \text{ 中的有效个数}}{8} = \frac{2}{8}$$

$$\text{不确定值} = \frac{\text{向量1中0的有效个数} - (\text{向量1} \cup \text{向量2}) \text{ 中0的有效个数}}{\text{向量1的有效位数}} = \frac{(\text{11010100}) \text{ 中0的有效个数} - (\text{11010100} \cup \text{11011101}) \text{ 中0的有效个数}}{8} = \frac{2}{8}$$

3.3.2 全局信任值的计算

全局信任值是通过全局信任表中的数据来计算的，全局信任表的构建与本地信任表类似，它比本地信任表要多存储 3 个变量： s_1, s_2, s_3 ，它们分别表示交易总次数，交易成功的次数，交易失败的次数。以 3.3.1 节中的假设为例，计算方法如下：

$$\text{信任值} = \alpha \times \frac{\text{向量1中的有效个数}}{\text{向量1的有效位数}} + \beta \times \frac{s_2}{s_1}$$

$$\text{不信任值} = \alpha \times \frac{(\text{向量1} \cup \text{向量2}) \text{ 中的有效个数} - \text{向量1中的有效个数}}{\text{向量1的有效位数}} - \beta \times \frac{s_3}{s_1}$$

$$\text{不确定值} = \alpha \times \frac{\text{向量1中0的有效个数} - (\text{向量1} \cup \text{向量2}) \text{ 中0的有效个数}}{\text{向量1的有效位数}} - \beta \times \frac{s_1 - s_2 - s_3}{s_1}$$

4 实验结果

本文对 BGTR 模型进行了模拟估算实验，为了验证模型的可靠性和可行性，本文用 JBuilder9.0 和 JXTA2.2 构建了模拟实验平台，设置了与实际应用相近的实验场景，进行了相应的模拟。

图 4 的是风险指数对交易成功率的影响，本实验共模拟了 20 个节点，并将节点初始全局信任值设为 [0.6, 0.7]，部分直接信任值设为 [0.6, 1.0]。共模拟了 3 024 次交易，实验结果显示，风险指数越高，不确定性因素作为积极因素考虑的越

多，交易成功率则越低。

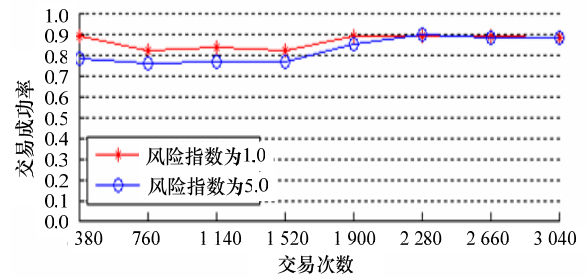


图 4 风险指数影响下的交易成功率比较

图 5 是本模型的抗干扰有效性测试结果，在存在一定的恶意推荐的情况下，随着网络中恶意节点数量的增加，当恶意节点比例达到 60% 时，BGTR 模型仍然可以达到比较高的交易成功率。当恶意节点大于 60% 时，虽然交易成功率变低，但交易次数也急速变少，P2P 网络中进行的交易很少，因此，交易失败的次数也很少。

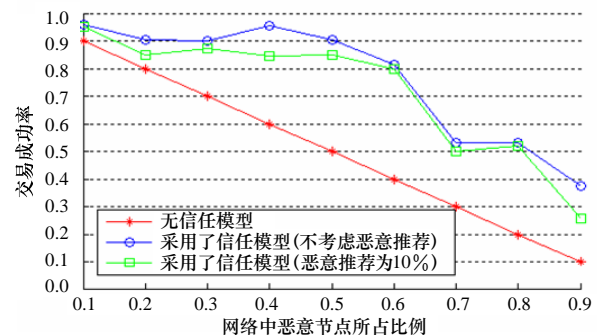


图 5 2种模型同等条件下的交易成功率

5 结束语

本文采用信任和推荐 2 种机制，构建了一个在 P2P 网络中的安全信任模型——BGTR，并在模型中加入了大量的实时因素，使得模型更能适应节点行为的动态变化。实验证明，此模型在交易成功率方面也优于之前的其他模型。

参考文献

- [1] Christin N, Weigend A S, Chuang J. Content Availability, Pollution and Poisoning in File Sharing Peer-to-Peer Networks[C]// Proceeding of the 6th ACM Conference on Electronic Commerce. [S. l.]: ACM Press, 2005: 68-77
- [2] 张春瑞, 徐 格, 王开云, 等. 基于信任向量的 P2P 网络信任管理模型[J]. 清华大学学报: 自然科学版, 2007, 47(7): 1125-1128.
- [3] Singh A, Liu Ling. TrustMe: Anonymous Management of Trust Relationships in Decentralized P2P Systems[C]//Proc. of P2P'03 Conference. Washington D. C., USA: IEEE Computer Society Press, 2003: 142-149.
- [4] Aberer K, Despotovic Z. Managing Trust in a Peer-to-Peer Information and Knowledge Management[C]//Proc. of P2P'01 Conference. New York, USA: ACM Press, 2001: 310-317.

编辑 索书志