基于网格环境中的量化评估信任模型

农 毅 古天龙
(桂林电子科技大学计算机系 桂林 541004)

摘 要 网格计算系统是由地理上分布的,异构的计算机和资源组成,它是一个分布式的高性能计算机环境,通过网络连接,用户可以透明地共享这些资源。但其本身具有动态性和不确定性,给区域间网格实体的合作带来一系列安全问题。所以网格环境中的信任问题,成为当前网格研究的一个热点。在本文中,把信任划分为身份信任和行为信任,提出了一种新型的量化评估信任模型,来处理网格环境中实体之间的信任关系,从而更科学、有效地解决网格环境中存在的安全问题。

关键词 网格,信任,信任度,信任模型

A Quantification-evaluating Trust Model Based on Grid Environment

NONG Yi GU Tian-Long
(Department of Computer Science and Technology, Guilin University of Electronic Technolog, Guilin 541004)

Abstract A grid computing system is composed of heterogeneous geographically-distrubuted computers and resources. And it is a high performance distribute environment. Users can share these resources pellucidly by the network. Because of the dynamics and uncertainty of grid environment, the collaboration of grid-entity is facing to a series of problem of security. So the trust issue of grid system becoms a hotspot in current study of grid. In this paper, we carve up the trust to identity-trust and behavior-trust, develop a new trust evaluating model to scale the relationg of the trust of grid entities, by which we can resolve security problems exist in grid environment more sciencely and effectively.

Keywords Grid, Trust, Confidence- vote, Trust model

1 引言

在当今结构复杂的社会有机体中,信任已经逐渐成为人类组织和个体存在的条件,没有信任,就难以进行真正具有创造性的合作,信任是合作的动力与源泉。

随着网格计算系统的发展,网格环境中的动态性和不确定性,使得信任问题已经成为安全保障的一个重要因素。当网格实体间进行紧密的合作时,为了能够可靠地共享资源,我们需要确定它们之间的信任关系。信任可以划分为实体身份的信任和实体行为的信任[1]。传统的安全机制有加密技术、访问控制等等,它们被用来提供授权和认证,解决了身份信任的问题。但它们并不能处理行为信任,不能保证提供你所需要的服务(例如不能保证按你所需要的方式进行授权)[2]。本文提出了一种新型的信任模型,在解决身份信任的基础上,采用数学的方法把实体间的行为量化,从而科学地评估实体之间的信任关系,帮助网格实体进行信任抉择,有效地解决网格环境中存在的安全问题。

本文第2节简单介绍了涉及信任问题的一些基本概念和相关定义;第3节具体描述了我们提出的信任模型,并给出了实现的方法;第4节对该模型进行实验模拟验证,分析实验结果;最后是结论部分。

2 信任问题研究

2.1 声誉,信任和信任度

关于声誉和信任的研究已经很多[3~5],它们是相关联的两个概念,都可以用来评价某个实体的信任度。在进行讨论之前,我们需要清楚相关的定义。

在本文中提出的信任模型中给出如下定义:

声誉(Reputation):是一般所说的一个人或物的特征或身份,它是一个全局的概念,是一定范围内的所有成员对一方的信任综合评价。在模型中,假定一定范围的网格实体集合为{P1,P2,P3,P4,…},如果P1对P2的信任度为Trust′,则把其它所有实体对P2的信任度集合以Trust′作为均值,求出方差M作为P2的声誉评估值。

信任(Trust):指特定条件下一方对另一方的能力、诚实和可靠性的主观的概率估计。信任可以基于直接和间接的经验来获得。在模型中,我们用信任度来表示信任等级的高低,用声誉评估值来确定可信任程度的偏差,信任度和声誉评估值随实体的行为而动态变化,信任度越大,声誉评估值越小,则P1对P2的可信任程度越高。

信任度(Trust′):指信任者进行某种决策时对被信任者的可信任程度,是一种对信任的量化估计。信任的程度可以从完全不信任到完全信任,对应于信任度范围从0到1,即信任度的取值为连续值,可以取0到1之间的任一值。

2.2 信任关系

网格实体集合{P1,P2,P3,P4,…},任意两个实体P_i和P_j的信任关系可以分为两类:直接信任和推荐信任[6]。在该模型中,我们给出以下定义:

直接信任:任意两个实体P_i和P_j之间如果有过直接的交易,则P_i和P_j相互建立了一种直接信任关系,根据双方的交易情况得出直接信任度。

推荐信任:任意两个实体P_i和P_j之间如果没有进行过

直接的交易，而是根据其他所有实体 Pk（$k=1,2,3,\cdots$，且 k 不等于 i 和 j）的推荐建立的一种信任关系，则 Pi 和 Pj 的信任度是根据其他实体的评估得出的结果。

两者的区别，我们可以从图 1，图 2 中可以看出。

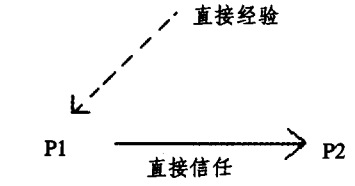

图 1　直接信任

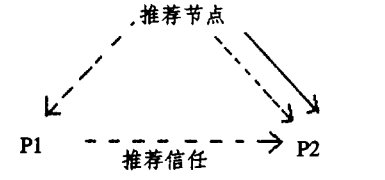

图 2　推荐信任

2.3　信任的特性

在本文介绍的模型中，网格实体集合{P1，P2，P3，P4，…}，任意两个实体 Pi 和 Pj 之间的信任具有以下特性：

· 主观性。不同的个体对同一事物的看法会受个体喜好等因素影响而有所不同。

· 有条件传递性。如 P1 信任 P2，P2 信任 P3，并不表示 P1 就信任 P3。但是通过 P2 的推荐，P1 到 P3 之间可以建立信任关系。但是 P1 对 P3 的信任程度还依赖于对推荐来源 P2 的信任程度。

· 动态性。直接交互和推荐都可以改变对一个实体的可信性的评估。信任和声誉都会随着进一步的经历动态地增加或减少，且不同的背景会使得一个实体对另一实体有不同的信任评价，可以是环境不同，也可以是时间不同等等。

· 传播性。信任可以通过推荐进行扩散，传递推荐的机制就是声誉机制，从而可以辅助其他实体进行决策[7]。

2.4　信任和安全的关系

通常情况下，安全机制是通过阻止恶意的实体来提供安全保障的。传统的安全机制通过加密技术、访问控制等等来保证整个系统的安全。然而，在网格环境，实体之间要进行紧密的合作和实现资源的透明共享，却要从提供资源共享者那里保护自己，信任问题和安全问题变得更加复杂。在这种情况下，传统的安全机制已不能解决此类安全隐患的威胁。研究证明，信任和声誉系统却能够解决此类问题。两者的区别首先被 Rasmussen 和 Jansson 提出来，他们使用术语“硬安全”来表示传统的的机制如认证和访问控制，用“软安全”来表示普通意义上的社会控制机制如信任和声誉[8]。

3　信任模型

网格是通过网络向用户提供可随时按需利用计算资源和信息资源的环境。这一环境中，地理上广域分布于网络上的多种多样的计算资源和信息资源（个人计算机、工作站、集群、科学仪器等），可以实现透明的共享。我们把连接到这个环境中的资源或用户称为一个网格实体，用 Pi 表示。当网格中任意两个实体要进行交易时，必须确定它们之间的信任关系，得到相互间的可信任程度，以作出安全的信任决策。

3.1　身份信任

网格环境下的身份验证较为复杂，主要是要考虑到验证身份的双方一般都是通过网络而非直接交互。同时大量的黑客随时随地都可以尝试向网络渗透，截获合法用户口令并冒名顶替以合法身份入网。在本文提出的模型中，我们采用高强度的密码认证协议 RFC1760，基于一次性口令思想开发的身份认证系统 S/KEY 来进行身份认证[9]。

3.1.1　信任连接

当一台 pc 客户机要连接到网格环境中，首先要通过服务器的身份认证。S/KEY 的认证过程分为以下几个步骤。

①向服务器发出连接请求命令 conect。

②服务器返回应答，发送两个参数：seed 和 seq。

③pc 客户输入口令，将口令与 seed 连接，使用 MD5 算法做 seq 次 Hash 计算，产生一次性口令，发送到服务器。

④服务器收到客户端发送过来的口令，进行一次 Hash 计算，并从存储文件中取出该用户上次登录的一次性口令做比较，如果匹配，则通过身份验证，再把收到的一次性口令覆盖掉原先的口令，同时 $seq'=seq-1$。

⑤当下次该用户发送连接请求，服务器将返回应答参数 seed 和 seq'，则重复③④完成下一次的身份验证。

3.1.2　S/KEY 身份认证的优点

客户机向服务器发送的口令是通过秘密口令和 seed 经过 MD5 散列算法得到的密文，由于散列函数的非可逆性，恶意的实体不可能破解网络上截获的密文，当然也不可能破解通过攻击服务器得到的密文。而在客户机每次成功登录后，服务器端都会自动更新存储在文件中的 seq 值（$seq'=seq-1$），使得客户机在下一次连接生成的发送给服务器的口令密文与上一次不相同，进一步增加身份验证的强度和力度。

3.2　行为信任

经过 3.1 节的讨论，连接进入网格环境的客户机已经通过身份上的安全认证，可以透明地共享自身和其它的网格实体资源。在这种情况下，仍然可能存在一些受病毒感染的、有目的性的、或者恶意的、有欺诈行为的网格实体，那么在网格实体进行共享或者交易的时候，要不要进行，安不安全的问题就显得尤为重要。在本文提出的模型中，我们用实体间的信任度和声誉评估值，来帮助作出信任决策。

3.2.1　信任度的格式

在传统的信任模型中，信任度通常采用一个十进制的数值来表示，当两个实体间发生一次真实的交易，信任度加 1；如果发生的是一次欺诈的交易，信任度减 1。在日常经验中我们可以发现，就算两个实体之间一直进行的是真实交易，但只要发生一次欺诈交易，对作出欺诈行为的实体，主观上对它的信任度也会大打折扣。从这个意义上讲，以上信任度的评估方法可行，但是存在一定的不足，不能把现实中人脑的主观意识和经验抉择很客观地表现出来。

在本文的模型中，我们采用新的格式——通过一个定长为 L 位的二进制向量（L 的值由网格区域的网格实体数量决定，可以采用 4，8，16，32，64，…，数量越大，进行共享或交易的概率和频率越大，L 的值也应越大。本文取 $L=8$ 作为讨论）来表示信任度，标志为 Trust。在这 L 位二进制数据中，“1”表示一次诚实的交易，“0”表示一次不诚实的交易。网格区域的网格实体集合{P1，P2，P3，P4，…}，如果 Pi 请求 Pj 的

服务，当 Pj 提供了一次真实的服务，则在 Pi 对 Pj 的信任度向量 Trust 的左侧插入“1”，相应地其他位数右移；如果 Pj 提供的是一次欺诈服务，则在 Pi 对 Pj 的信任度向量 Trust 的左侧插入“0”，相应地其他位数右移。同时用一个整数 t 来记录两者发生交易或服务的次数，不管是真实的还是欺诈的，只要发生一次，t 增加 1。

举例说明如下：假设 Pi 对 Pj 的信任度为 Trust＝10010000，到目前为止发生交易或服务的次数为 t=4。那么，如果 Pj 为 Pi 提供一次真实服务，那么 Trust＝11001000，t=5；如果 Pj 对 Pi 提供的是一次欺诈服务，那么 Trust＝01001000，t=5。

3.2.2 信任度的动态存储

在传统的信任模型中，P1 作为一个信任实体，通常是把它对实体 P2 的信任度存储在本机文件中，或者统一被存储在一个服务器中，这给信任度的查找带来很大的方便。但是在另一方面，却容易给黑客或者带恶意性的其它实体所攻击，出于某些方面的原因篡改信任度。

例如，P1 在以往和 P2 的数次交易中，P2 提供的都是欺诈服务，则 P1 对 P2 的信任度很低，P2 其实是一个恶意实体。而这个信任度 Trust 被存储在 P1 所在的客户机中，或者存储在一个服务器中。那么当 P1 在下次与 P2 要发生交易或者服务请求，P2 为了能够得到 P1 的信任，它会想方设法攻击存储 Trust 所在的 PC 机或服务器，试图篡改 Trust 的值，以达到得到 P1 的信任的目的。总之，只要让一个带有恶意性或欺诈性的实体轻易知道对它的信任度 Trust 存储的地方，安全性真实性就受到威胁，易受攻击。

另一种传统的存储方式为了解决上面提出的问题，采用分散存储方式，即信任度 Trust 的信任不存储在服务器中，而是存储在进行交易的各个 pc 机中，如 P1 对 P2 的信任度信任存储在 P1 对应的 pc1 中。这样大大减少了上述服务器存储的威胁。但同时带来相应的问题：当网格环境的 pc 机下线后，将不能把存储在它上面的信任度信息发送给需要该信息的 pc 机。这样，当大量 pc 机离开网格环境后，将使信任评估结果与真实情况出现很大差异。

基于以上两种情况，在本文提出的模型中，信任度的存储与传统的方式不同，我们采用动态的存储方式，把信任度经过加密后，动态、随机地存储在网格环境的 PC 机中。

假定网格环境中的实体集合{P1，P2，P3，P4，…}，它们对应所在的 PC 机集合为{pc1，pc2，pc3，pc4，…}，P1 对 P2 的信任度为 Trust，则对 Trust 的存储经过以下三个步骤：

①随机选定目标。在 PC 机集合{pc1，pc2，pc3，pc4，…}中随机搜索到一台当前在线的 PC 机，假定为 pc3。它是一个随机的过程，其它恶意实体不可能知道选定的哪个 pc 机，而且它保证选中的 pc 机是当前在线的。

②双重数据加密。

• pc1 选定 pc3 后，向 pc3 发出请求命令 comand，并附带参数 PK1(pc1 的公开钥)；

• pc3 收到请求命令后，记录下 PK1，并向 pc1 作出回应，并附带参数 PK3(pc3 的公开钥)。

• pc1 在收到 pc3 的回应后，记录下 PK3(如果 pc1 收不到 pc3 的回应，则回到步骤①)；接着用自己的私人秘钥 SK1 对信任度明文 Trust 进行加密，过程表示为 c＝Esk1[Trsut]；再用 pc3 提供的公开钥 PK3 对加密后的密文 c 再进行一次加密，过程表示为 c′＝Epk3[c]，即 c′＝Epk3[Esk1[Trsut]]。最后向 pc3 发送数据，数据格式为：

id1(实体 P1 的标志)	id2(实体 P2 的标志)	c′(双重加密后的密文)

• pc3 收到 pc1 发送过来的数据后，存储到相应的文件目录中。

③存储转移。当 pc3 要下线离开网格环境，为了保证存储在本机上的信任度信息能够及时提供给需要的网格实体，需要转移存储到当前在线的其它 PC 机。首先，pc3 用记录下的 pc1 的公钥 PK1 和自己的秘钥 SK3 对第三个数据段 c′进行解密，过程表示为 Trust＝Dpk1[Dsk3[c′]]。第一和第二字段的值保持不变，重复①②步骤，完成数据存储的转移。

3.2.3 信任度的查找

网格实体集合{P1，P2，P3，P4，…}，它们对应所在的 PC 机集合为{pc1，pc2，pc3，pc4，…}，任意实体 Pi 要与另外的任意一个实体 Pj 发生交易或合作，Pi 要从整个网格环境中查找收集到它本身和其它实体对 Pj 的信任度，以便作下一步的信任量化评估。它的查找过程分为以下三个步骤：

①发送请求。pc$i$ 向网格环境中所有在线的 PC 机发送需要查找对 Pj 信任度的请求命令 comand，并附带参数 IDi(Pi 的标志)，IDj(Pj 的标志)，PKi(pc$i$ 的公开钥)。

②请求回应。当 pc$k$(k＝1，2，3，4，…)收到请求，根据 IDi 与 IDj 与本机存储的信任度数据做比较，如果符合要求，则根据 3.2.3 中③步骤的方法进行数据存储转移，转移目标为 pc$i$。

③pc$i$ 在收到其它 PC 机回应的信任度数据后，对 c' 进行数据解密，存储在一个自定义链表中。

3.2.4 信任度的标量化

为了计算实体间的信任度和声誉评估值，我们需要将二进制的信任度格式标量化，得到十进制的信任度格式 Trust′，用于参与数学计算。在该模型中，信任度的标量化分为以下两种情况：

①当交易次数 t 小于或等于 L 时，Trust′＝(前 t 位可信向量所对应的二制值)/2^t。

②当交易次数 t 大于 L 时，Trust′＝(L 位可信向量所对应的二制值)/2^L。

两种情况下信任度标量化的举例如下：

当实体 Pi 对 Pj 的信任度向量是 Trust＝10010000(L＝8)，t=4 时，

$$\text{Trust}'=\frac{(1001)_2}{2^4}=0.5625;$$

当实体 Pi 对 Pj 的信任度向量是 Trust＝10011010(L＝8)，t=9 时，

$$\text{Trust}'=\frac{(10011010)_2}{2^3}=0.6015。$$

经过以上两种情况下对向量信任度的标量化，可以得到一个从 0～1 之间的连续小数，它就是某一实体对另一实体的标量信任度。

3.2.5 信任量化评估

网格实体 Pi 要对 Pj 的信任度和声誉进行量化的评估，经过 3.2.3 节对它信任度的查找，并经过 3.2.4 对所有向量信任度的标量化，我们采用数学的方法进行处理，分为以下两种情况。在讨论之前，我们进行相应的一些约定：假定 Pi 对 Pj 的标量信任度为 Trust1′，其它实体对 Pj 的标量信任度分别为 Trust2′，Trust3′，Trust4′，…。

①如果 Pi 对 Pj 有直接信任度 Trust1′。

·把 Trust1′作为信任度均值衡量。

·计算 Pj 的声誉评估值 M。其中 M 为集合{Trust1′, Trust2′, Trust3′,Trust4′,…}以 Trust1′ 为均值的方差。计算公式为：

$$M=\sqrt{\sum(\text{Trustk}'-\text{Trust1}')^2}$$

其中 $k=1,2,3,4,\cdots$。

·先根据 Trust1′的值，可以确定 Pi 对 Pj 的信任度等级（等级可以分为完全信任，信任，不大信任，完全不信任等，可以根据需要由标量信任度的取值范围而定）。再根据 M 的值评估 Pi 对 Pj 的信任度 Trust1′的支持度，M 为其它实体对 Pi 的信任度在 Trust1′的约束下的客观评估值，即信任 Pj 的程度为 Trust1′ 的偏差。根据经验，如果 M 的值越大，即偏差越大，那么直接经验下 Pi 对 Pj 的信任度为 Trust1′ 越值得怀疑；如果 M 的值越小，即偏差越小，那么直接经验下 Pi 对 Pj 的信任度为 Trust1′ 越真实。最后我们可以根据 M 的取值范围，对信任度的等级适当进行降级。

②如果 Pi 对 Pj 没有直接信任度 Trust1′。

·计算集合{Trust2′, Trust3′, Trust4′,…}的均值作为 Trust1′的值。公式为：

$$\text{Trust1}'=\frac{\sum_{k=2}^{n},\text{Trustk}'}{n}$$

其中 $k=2,3,4,5,\cdots$。

·执行①中第二，第三步骤。

4 实验及结果分析

在本文中，提出的新型量化评估模型，保证了网格环境中对身份信任和行为信任的处理。基于身份信任的基础，在我们的行为信任评估过程中，通过采用新的信任度格式，是信任的衡量更符合社会经验，并通过把双重加密后的信任度随机的存储在网格环境中，因为它具有很高的隐蔽性，使得在进行信任评估的过程中，极大地避免了受到一些恶意实体的攻击。在最后的信任评估当中，把现实经验中的信任度和声誉都考虑到模型中，不完全取决与直接信任度的主观经验，而是把所有的实体对某一实体的经验进行量化，使得模型评估出的结果更具有科学性和准确性。为了验证该模型的准确性和科学性，我们进行了仿真实验。在我们的实验中，假定对信任度的查找过程是一个理想的状态，即能把存储有需要的信任度信息都能查找到。在模拟环境中，每次随机地选取一定比例的节点作为恶意实体，提供错误的评估信息。

实验表明，在传统的计算方法下，信任度的评估很容易受到恶意实体的干扰，评估的准确度波动很大。因为任何一个恶意实体都可能攻击其它存储有信任度信息的 PC 机，对信息进行篡改，提交一个不真实的信息来影响整个评估系统的准确性。如图 3 所示，本文所采用的模型和计算方法不容易受这种恶意实体的干扰和影响，它能更准确、更科学地反映实体之间的信任关系。

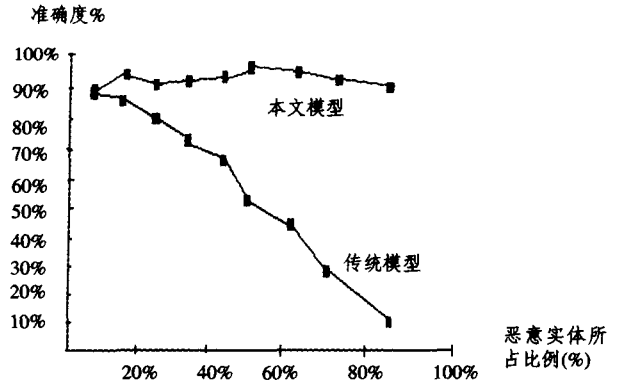

图 3 实验结果分析图

结论 本文从网格环境中的信任安全考虑，提出了一个新型的量化评估信任模型，包含了身份信任和行为信任。身份信任采用了基于一次性口令思想开发的身份认证系统 S/KEY 来进行身份鉴定，保证连接进入网格环境的实体都是合法用户。行为信任中把实体间的信任量化为对信任度和声誉评估，信任度格式有别于传统的风格，采用向量表示法，使其表达的意义和结果更符合社会经验的真实情况。为了保证信任度信息安全，我们采用随机存储机制，系统模型具有极大的隐蔽性，使环境中一些恶意的实体不可能攻击到。在最难衡量的行为信任中，我们采用数学的方法，把所有信任度量化为一个数据集合，通过计算把均值作为直接信任经验，同时计算出方差作为直接信任经验的偏差，共同衡量实体间的信任关系。最后，实验结果表明，该模型能够科学、有效地解决网格环境中的安全问题和信任问题。

参 考 文 献

1 Azzedin F, Maheswaran M. Towards Trust-Aware Resource Management in Grid Computing Systems. In: Proceedings of the 2nd IEEE/ACM International Symposium on Cluster Computing and the Grid, 2002

2 Yu Bin, Singh M P. A Social Mechanism of Reputation Management in Electronic Communities. In: Proceedings of the 4th International Workshop on Cooperative Information Agents, 2000. 154～165

3 Marsh S. Formalizing Trust as a Computational Concert: [Ph. D. Thesis]. University of Strirling, 1994

4 Freeman L C. Centrality on Social Networks. Social Networks, 1979, 1: 215～239

5 Marsden P V, Lin N. editors. Social Structure and Network Analysis. Beverly Hills: Sage Publications, 1982

6 Borcherding B T. Valuation of trust in open networks[C]. In: Proccedings of the European Symposium on Research in security, Brighton: Springer-Verglag, 1999. 59～63

7 张书钦，芦东昕，杨永田. P2P 文件共享网络中信任管理系统的设计. 计算机工程，2005，31(18)：160～162

8 Rasmusson L, Janssen S. Simulated Social Control for Secure Internet Commerce. In: Catherine Meadows, ed. Proceedings of the 1996 New Security Paradigms Workshop. ACM, 1996

9 张世永. 网络安全原理与应用. 科学出版社 2003. 171～173