自治网络中信任信誉模型的安全现状研究

孙玉星[1,2] 黄松华[1] 黄 皓[1] 谢 立[1]
(南京大学计算机科学与技术系计算机软件新技术国家重点实验室 南京 210093)[1]
(南京审计学院信息科学学院 南京 211815)[2]

摘 要 随着P2P网络、Ad hoc、无线传感器网络的深入研究,信任和信誉成为保障这类自治网络安全的一个重要手段。虽然信任信誉系统在自治网络中起到了重要的作用,但其采用了间接推荐等技术,给信任信誉带来很多安全问题。介绍了信任信誉模型的相关概念,总结了目前对信任信誉模型的新攻击手段,并针对这些攻击,比较分析了在自治网络环境中具备一定防御能力的典型信任信誉模型的各自防御方法、防御效果以及性能情况。最后,在分析了现有研究存在的主要问题的基础上,展望了今后提高信任信誉模型安全性研究的主要方向。

关键词 自治网络,信任模型,信誉模型,网络安全
中图法分类号 TP393 **文献标识码** A

Research on Security of Trust and Reputation Model in Autonomous Networks

SUN Yu-xing[1,2] HUANG Song-hua[1] HUANG Hao[1] XIE Li[1]
(State Key Laboratory for Novel Software Technology, Department of Computer Science and Technology, Nanjing University, Nanjing 210093, China)[1]
(School of Information Science, Nanjing Audit University, Nanjing 210031, China)[2]

Abstract With the in-depth research of peer-to-peer networks, Ad hoc networks, wireless networks etc, trust and reputation become important means to the protection of Autonomous Networks. Although the trust and reputation managements have played an important role, indirect recommendation techniques bring a lot of security issues in the trust and reputation systems. After the concepts of trust and reputation were presented, some novel attacks related were given. And then several typical trust/reputation models with defense capabilities for the self-government networks were analyzed and compared in security and performance after they were described briefly. Finally, the current problems and also challenges in the future in security of trust and reputation model were presented.

Keywords Autonomous networks, Trust model, Reputation model, Network security

随着网络技术和应用的快速发展,传感器网络和普适计算等技术应运而生,这些变化给网络管理和控制带来新的挑战。网络的移动性和动态性特征、大量的数据以及信息交流均需要新的方法解决管理问题。传统的集中基于服务器的管理,不能再满足下一代网络的需要,人们开始提出新的网络基础设施和管理概念。例如,旨在利用任何基础设施提供无线网络服务的移动Ad Hoc网络[1](MANETs)和可以实现数以百万的网络用户共享大量数据的点对点(P2P)网络[2,3]等等,它们都有共同特征:分布性和自组织性,通常被称为自治网络。

随着各种自治网络的出现,Internet的安全问题受到越来越严重的威胁。传统的提供网络安全做法一直是借用工具进行加密和验证。但是,基于加密的安全技术不能解决在自治网络中所遇到的新特征和新恶意行为,因为加密技术不能防止内部对手或错误节点提供有意的和无意的错误服务。因此,对信任/信誉管理逐渐成为保障自治网络安全的重要部分。很多研究者对信任/信誉系统作了各方面研究[4-8],例如采用信誉技术激励Peer增强协作、惩罚恶意行为等。这些研究工作表明,在自治的网络中使用信誉系统,大量地减少P2P网络、Ad hoc网络上的恶意行为。

虽然信任/信誉系统在自治网络安全中扮演了重要角色,但是由于信任/信誉中采用了间接信任传递手段,攻击者能采用各种方法阻碍信任机制建立正确的信任/信誉评价,从而达到破坏信任/信誉系统安全的目的。本文总结了攻击信任/信誉机制的新手段,并分析了现有的典型信任模型对各种新攻击手段的抵御能力,为以后研究的进一步开展打下基础。

本文组织如下:第1节主要介绍信任和信誉的基本概念以及常用模型的分类;第2节总结信任/信誉模型新威胁;第3节分析自治网络的特征对信任/信誉系统的安全带来的问题以及设计信任/信誉系统所需要满足的需求;第4节选取新

到稿日期:2008-05-23 本文研究得到国家自然科学基金项目(No. 60303023)资助。
孙玉星(1977—),讲师,博士研究生,研究领域为无线网络、网络安全,E-mail:xyusun2006@gmail.com;黄松华(1979—),博士研究生,研究领域为网络移动、网络安全;黄 皓(1957—),教授,博士生导师,研究领域为计算机安全与网络安全;谢 立(1942—),教授,博士生导师,研究领域为信息安全、分布并行计算。

的、典型的适用于自治网络的、具有一定攻击抵御功能的信任/信誉模型进行分类评述，主要分析它们各自在抵御攻击上的优缺点及其性能上的差异；第5节阐述了目前该领域还存在的问题和以后需要进一步研究的内容。

1 信任和信誉的概念

1.1 信任的概念

很多文献都认识到信任计算模型的重要性，其研究内容从专用应用模型到一般模型，但很多作者所提到的信任的定义是随着研究者各自的研究应用领域的不同而不同的。我们总结了常见的信任定义有以下两种。

定义1 信任[11]是一个实体评估另一个实体或团体将执行某种行为的一种特定的主观概率，这一评估是在实体可以观察到行为之前，是上下文相关的，并会影响它自己的行为。

定义2 关于某服务 X，一个实体 A 对实体 B 的信任是一种可测量的 A 对 B 在一定时期内，一定相关内容上（和服务 X 相关）的行为的信任[9]。

由以上两种定义可以看出，信任包含两方面的特征：主观性。不同的实体对于同一个目标实体可以有不同的信任，信任是一种基于很多因素或证据的个人的主观现象。上下文相关性。任何一个信任都是和一定的内容相联系的，例如对一个实体某项操作可信，并不等于某一个实体的推荐行为可信。

1.2 信誉的概念

定义3 信誉[13]是对一个人或一个事物的特性或行为的普遍认识。

信誉是根据社团中成员的推举和投票所产生的综合可测量的信任度。个人的主观信任可以由收集到的推荐和个人经验推断出来。信誉与信任度的概念是紧密联系的，计算信任度的一种重要方法就是基于信誉计算结果，但是两者之间也有很大的区别[13]，主要在以下两个方面：第一，信任系统所产生的信任是信任主体对信任对象的信任度的主观评价，而信誉系统是根据社团成员的投票而产生的实体的公共信誉值。第二，在信任系统中，信任的传递性是一个显性的手段，而在信誉系统中可以隐性地参考信任传递，甚至不利用信任的传递特性。

1.3 信任信誉的计算

信任信誉的计算根据信息源的不同可以归结为两类：一种是直接信任；另一种是间接信任。直接信任是实体对其他实体的可信度的一种独立信任，这种信任是建立在与其他实体的直接交互经验的基础上，通过对与其它实体交互过程，观察对方行为优劣从而建立对其他实体信任度的评价。间接信任是一些实体对其他实体信任度的信任，这种信任的信息源是来自团体中其他成员的信息，这些信息可以是它们自己的直接信任，也可以是它们从其他成员那里收集的信息。一般我们把从其他成员那里收集到的信息称为推荐。正因为在信任中使用了间接的信任信息，所以信任具有传递性。

信任计算按其采用的定量研究方法[14]的不同可以分为以下两类：

概率论模型：信任计算中采用了统计学中概率论原理。采用这种方法的好处是信任计算有了良好的数学工具做基础，并且有丰富的概率论方法可以运用，其中使用最多的是贝叶斯方法。在以概率论为基础的信任计算中引入了信仰理论(Belief Theory)，其基本思想是：所有可能结果的概率和不一定为1，剩下的那部分概率可以解释为不确定性。

流模型：系统通过环或者任意长的链的传递迭代计算信任和信誉度，这种模型称为流模型。一些流模型假设团体中有一种固定的信任/信誉权值。常见的采用流模型的系统有：Google's PageRank[17]，Appleseed algorithm[18]和 Advogato's reputation scheme[19]。

2 自治网络环境的特点

信任反映了网络活动中参与实体的一组关系。在传统网络中，例如 Internet，信任证据集中在控制服务中心，例如第三方信任中心(TTPs)和认证中心。这些服务是一直被信任且有效的。但是在自治网络，没有固定基础设施和集中的控制服务器，在这些网络中信任证据是同等的，即来自组成网络的各个实体。本文主要讨论在自治网络中信任和信誉系统的安全问题，因此需要了解一下自治网络中实现信任和信誉系统的特点。

- 不确定性和不完整性：信任信誉证据是由组成网络的各个实体提供的，这些证据是不确定的，也不一定全面，甚至可能是不正确的。
- 本地性：所有的信任、信誉信息的交换都是通过各自与邻居实体交互完成，具有本地特征。因为具备本地性特征，这可以节省网络的资源（能量、带宽、计算能力）。
- 分布式计算：信任和信誉的计算以分布式方式执行。采用了分布式的计算方法可以避免单点失效，并且可以适应拓扑结构以及关系的变化。

由于自治网络中使用的信任信誉系统具备分布式特征，因此管理工作难度加大，比如证据的收集、规则的制订、估计的规范性等等，而且自治网络中信任信誉系统的安全问题也比常见的集中式信任信誉系统复杂。下面我们将分析自治网络中信任信誉系统的新攻击行为。

3 信任/信誉模型的新威胁

信任信誉系统虽然可以有效提高自治网络性能，并检测出恶意实体，但是信任信誉系统本身也是攻击者攻击的对象。除了一些常见的攻击手段，本文主要讨论信任信誉系统由于自身所采用的方法而引入的一些新的安全问题。

- 错误推荐攻击：因为现有的大多数信任信誉系统中都将他人推荐作为计算信任信誉的参考量之一，所以恶意实体都可以通过不诚实的推荐，暗害好实体或提高恶意实体的信任度。错误推荐攻击是一种最直接攻击手段。在信任系统中，这种攻击常被称为 Bad mouthing 攻击；在信誉系统中，这种攻击常被称为 Unfair rating 攻击[21]。因为这两种攻击本质上都是通过直接提供错误推荐从而影响正确的信誉值或信任度，所以本文归纳成错误推荐攻击。
- 叛国者攻击：这种攻击是恶意实体通过行为交替好坏，希望在保持高信任度的前提下实施攻击。一些恶意实体可以通过一段时间的良好行为建立高信誉度或信任度以后实施攻击。这种攻击利用了信任的动态特性，通过时间域上的不一致行为达到攻击效果。
- 偏见攻击：攻击者的不一致行为可以表现在时间域上，也可以表现在用户域上。恶意实体通过对不同实体提供不同

效果的服务来削弱对好节点的评价，这种攻击也称为冲突行为攻击。例如，服务者可以对一组用户始终表现好的行为，而对另一组用户始终表现坏的行为，那么这两组用户将对恶意实体行为产生冲突的认识，结果将导致两组用户间的相互不信任。

• 联合攻击：多个恶意实体可以联合起来给信任和信誉系统造成比单一攻击者更大的伤害，例如最常见的联合错误推荐攻击，即在信任或信誉系统中有多个说谎者，为系统提供错误推荐信息。如果说谎者数量超过一定限度，将导致信任和信誉系统失效。

• Sybil 攻击：传感器、Ad hoc、P2P 等自治网络中，存在一种称为 Sybil 攻击的有害攻击。它是一种通过节点非法宣称多个身份而实现攻击，一般称具有多余身份的设备为 Sybil 节点。已有很多文献分析了 Sybil 攻击对 Ad hoc 安全路由以及 P2P 文件共享的影响[20]。同样，Sybil 攻击也可以方便实现对信任信誉系统的攻击[22]。例如，一个用户可以创造多个虚假用户，这些用户相互连接并与攻击源点连接，为攻击源点提供高的信任信誉推荐，从而达到提高攻击源点信任度的目的。即使信任信誉系统没有受到直接攻击，恶意实体可以利用其 Sybil 节点代替恶意实体的恶意行为受过，即信任信誉系统仅将 Sybil 节点的信任度降低而并没有惩治到攻击源。

• Newcomer 攻击也称为洗白攻击[21]。在很多匿名场合，实体可以通过有意的离开再加入系统来消除以前身份的坏的信任或信誉度，即通过简单的撤离再加入轻易的消除以前身份的历史。很显然，Newcomer 攻击降低了信任和信誉系统的效率。

Sybil 攻击已在自治网络安全研究中引起高度重视，解决 Sybil 攻击的主要方法是提供保证身份与实体唯一对应关系的手段[20]。如果能很好解决保持身份与实体对应关系，也可以解决 Newcomer 攻击。这两种攻击虽然也会对信任信誉系统产生影响，但是如果解决身份与实体唯一对应关系以及该关系的保持问题就可以解决上述两种攻击，所以本文没有将这两种攻击作为参考标准。

4 典型模型及其评论

本文主要通过对自治网络中流模型和统计模型的信任/信誉系统的安全性进行分析，试图基本了解现有信任/信誉系统的安全现状，为以后的研究工作指明方向。

4.1 流模型

4.1.1 George's model(trust model based on semiring)

George[15,16]等人提出了一种基于半环(semiring)代数理论的信任模型。他们认为，信任推理问题就类似于在带权的有向图 $G(V,E)$(trust graph)上寻找最短路径的问题。图中用节点表示实体，有向边表示信任关系，然后使用半环代数理论计算两个节点之间的信任值，并进行信任评估。节点 i 到节点 j 的带权值的边表示了实体 i 对实体 j 的观点。权重函数定义为 $I(i,j):V\times V\to S$。S 是一个观念空间，由 trust 和 confidence 两个分量组成。trust 是一个信任估算值；confidence 是两个实体间经过多次交互后确立的信任估算的可靠性，代表了信任的质量。

本模型中最大的贡献是提出了采用半环代数理论抽象信任推理过程中所使用的计算方法。半环是一种代数结构$(S,\oplus,\otimes)$，其中 S 是一个集合，$\oplus,\otimes$ 是 S 上满足一系列属性的两个不同二元操作，$\otimes$ 操作可以用来求解一条路径上各个节点的推荐的合计，$\oplus$ 操作可用来求两个实体之间多条路径的合计。分别给 $\oplus,\otimes$ 定义不同的运算，半环可以体现出不同的运算特征，达到不同的运算效果。George's model 可以完成两个任务：第一，根据中间节点的信任度计算源端 A 应该赋予对目的点 B 的信任度；第二，找到节点 A 和节点 B 之间最值得信任路径。

文献中提到的一种距离半环所对应的 $\oplus,\otimes$ 操作如下：

$$(t_{ik},c_{ik})\otimes(t_{kj},c_{kj})\to\left(\frac{1}{\frac{1}{t_{ik}}+\frac{1}{t_{kj}}},c_{ik}c_{kj}\right) \tag{1}$$

$$(t_{ij}^{p1},c_{ij}^{p1})\oplus(t_{ij}^{p2},c_{ij}^{p2})\to\left(\frac{c_{ij}^{p1}+c_{ij}^{p2}}{\frac{c_{ij}^{p1}}{t_{ij}^{p1}}+\frac{c_{ij}^{p2}}{t_{ij}^{p2}}},c_{ij}^{p1}+c_{ij}^{p2}\right) \tag{2}$$

其中，数对(t,c)为观念空间的值。利用半环求最值得信任路径的算法是对 Dijkstra 算法的改进，具体参见文献[15]。

该模型主要采用了带权有向图表达实体间的信任关系，借助半环代数理论较好地抽象多种计算模型。在本系统中，推荐信任度可以使用多级信任链的方式，较全面地收集其他节点的信息。文章在模拟结果中，对该模型抵御攻击的能力做了一定的分析。作者将节点分为好节点和坏节点，好节点是根据预定规则如实调整它们的直接观点；而坏节点总给周围的坏节点赋最好观点(1,1)，给周围的好节点赋最坏的观点(0,1)。根据作者的模拟结果，模型具有较好的恶意节点的监测能力。

文章的实验模拟反映该模型具备较好的对错误推荐攻击的抵御能力，但是实验的前提是坏节点一定给出极端错误推荐，即对所有好节点采用最小化信任推荐值的策略，对所有坏节点采用最大化信任推荐值的策略，因此该系统对偏见攻击防御能力较弱。同样，该信任模型在计算时没有考虑到时间因素，无法体现信任的动态性特征，这为叛国者攻击提供了便利。对于联合攻击，如果联合策略是简单的，即单纯最大化周围坏节点信誉度，最小化周围好节点信誉度，那么该系统具有一定的抵御能力。但是对于复杂的联合攻击，该模型所能达到的效果有限。

4.2 统计学模型

4.2.1 Tao Jiang's model

Tao Jiang[25]等人提出了基于统计学的信任评估规则，采用马尔科夫链证明了该方法的收敛性，并分析了固定的网络结构下的一些特征。作者主要针对自治网络的一些特点提出了分布式计算信任评估的方法，该方法主要使用马尔科夫链分析本地一系列节点的信任评估值随时间的变化情况，总结信任评估值的随机分布，并在此基础上对模型判断结果的正确性进行定量的分析。

该系统使用了带权有向图 G 反映信任关系，图 G 只表现了在物理结构上的节点与节点之间的单跳关系，即邻接关系。图 G 中节点 i、节点 j 之间的有向边(i,j)表示实体 i 和实体 j 之间的直接信任关系；边上的权值表示 i 对 j 的信心程度，用 c_{ij} 表示，取值范围为$[-1,1]$。如果 $c_{ij}=1$，表示实体 i 对实体 j 的信任度有积极的信心；如果 $c_{ij}=0$，表示实体 i 对实体 j 的信任度完全不确定；如果 $c_{ij}=-1$，表示实体 i 对实体 j 的信任度是消极的信心。信任图如图 1 所示。

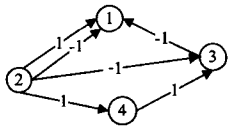

图 1　信任关系图

该系统的信任评价的基本运算涉及到 3 个基本量：$T=\{t_1,\cdots,t_N\}$反映了每个实体的真实信任度，$S=\{s_1,\cdots,s_N\}$反映了使用统计学方法得出的推断信任度，c_{ij} 也成为信心值。s_i 和 t_i 的取值为−1 或 1，1 代表好节点，−1 代表坏节点。信任度的估算采用了简单本地投票方式，类似于 reputation 系统信誉度计算方法，但是参与投票的不是系统中的所有节点，而是被估算信任度的目标节点的所有邻接节点。整个信任评估可以看成是随着时间不断演化的自动进程，其过程可以用以下等式表达：

$$s_i(k+1)=\begin{cases}1, & \text{当 } m_i(k)\geqslant\eta\\ -1 & m_i(k)<\eta\end{cases} \tag{3}$$

$$m_i(k)=\sum_{j\in N_i}\hat{c}_{ij}s_j(k) \tag{4}$$

其中$\hat{c}_{ji}=c_{ji}+ac_{ij}$，其主要作用是缓解投票节点和目标节点之间的信心冲突问题；η 为信任度阈值。

该模型采用了马尔科夫链分析所有节点的信任评估 S 所能达到的稳定状态的随机分布。马尔科夫链的状态分布概率如下：

$$\pi_S=\frac{e^{\beta U(S)}}{\sum_S e^{\beta U(S)}} \tag{5}$$

其中 $U(S)$为状态 S 所具备的能量，计算公式如下：

$$U(S)=\sum_{(i,j\in E)}(c_{ij}+c_{ji})s_is_j-\eta\sum_{i\in V}s_i \tag{6}$$

关于如何使用 S 的状态分布来定量评估信任评估的成功与否，已经超出本文要讨论的范畴，可以参见文献[25]。

该模型模拟阶段将攻击者可能采用的攻击模式分为 3 类：①独立式攻击。类似错误推荐攻击，这种攻击情况下攻击者不和其他攻击者串谋；②共谋式攻击。类似联合攻击，当然文章中提到的联合攻击是最简单的联合攻击，即对共谋者投最高信任票，对好节点投最低信任票；③任意攻击。这种攻击类似于叛国攻击和偏见攻击，没有固定的规律可循，攻击者随意发出不正确的投票评价。通过最后的模拟分析可以看出，该系统在面临独立式攻击和简单共谋攻击时都有较高的评估正确率，但是面临任意攻击时，系统的评估正确率随着攻击者数量的增加而急剧下降。

该系统还存在信任评估过于简单化的问题。对于实体的信任评价只采用 1 与−1 来区分，这大大降低了系统的通用性；而且文中用于模拟分析时采用的网络拓扑结构图是简单的二维格结构，即每个点都只有 4 个邻接节点，这与实际所使用的信任图有很大差异。作者将对复杂拓扑结构情况下该系统的性能做进一步定量研究。

4.2.2　Sonja Buchegger's model

Sonja Buchegger[7]等人专门针对信誉系统的欺骗问题提出了一种计算模型，这种模型基于传统的贝叶斯框架，在此基础上做出了一定的改进。下面分析 Sonja Buchegger's model 的主要思想和实际效果。

在该模型中，每个实体 i 上都需要保存关于其他实体 j 的三方面信息：R_{ij}，T_{ij} 和 F_{ij}。R_{ij} 表示实体 i 认为实体 j 在使用环境中行为可信的观点；T_{ij} 表示实体 i 认为实体 j 在信誉系统中的诚实度；F_{ij} 表示实体 i 对实体 j 的第一手观察信息。在该系统中实体与实体之间只传递第一手观察信息，即 F_{ij}。该系统采用贝叶斯方法来处理信誉投票信息，模型中都采用 Beta(α,β)分布为先验分布。该信誉系统具体的投票过程分为 3 步：

(1)当实体 i 得到关于实体 j 的第一手观察资料后，实体 i 就会更改本地保存的 F_{ij} 和 R_{ij}。对 F_{ij} 的更新如下：

$$\begin{aligned}\alpha&:=u\alpha+s\\ \beta&:=u\beta+(1-s)\end{aligned} \tag{7}$$

如果观测到目标的错误行为，$s=1$，否则 $s=0$。u 为时间折扣因子，降低历史观察信息对目前评价的影响。

(2)实体不时地发送自己的第一手信息给团体中的其他实体。例如实体 i 接收到实体 k 发来的关于实体 j 的 F_{kj}，如果 k 是可信实体，或者 F_{kj} 与 R_{ij} 类似，实体 i 就接受 F_{kj} 并修改 R_{ij}，否则 R_{ij} 保持不变。R_{ij} 的更新方法类似 F_{ij}。判断 F_{kj} 与 R_{ij} 是否相似的方法如下：

假设 $F_{kj}=(\alpha_F,\beta_F)$并且 $R_{ij}=(\alpha,\beta)$，如果$|E(Beta(\alpha_F,\beta_F))-E(Beta(\alpha,\beta))|\leqslant d$($d$ 为固定的正数，是差距阈值)，则 F_{kj} 与 R_{ij} 相似。

(3)系统执行过程中要不断地更新 T。当 F_{kj} 与 R_{ij} 类似，则提高 T_{ik}，否则就降低 T_{ik}。T_{ik} 的更新方法类似 F_{ij}。

该模型将外界传来的第一手资料和自己保存的信誉度相比较。如果偏差过大，就认为提供信息的一方是撒谎者，降低其可信度。通过模拟试验，可以看出在撒谎者较少的情况下，系统达到较好的抵御功能；当撒谎者较多时，模型对撒谎者的抵御能力将下降。其次，文中提到撒谎者虽然可以通过提供偏差度小于 d 的谎言来进行攻击，但是小幅度的谎言所达到的效果将随着时间流逝而减弱。再次，文中所分析的都是针对独立的说谎者，没有对联合攻击加以分析。从其工作的基本原理来看，如果有多个撒谎者作出协同的攻击，该系统将无法通过 F_{kj} 与 R_{ij} 是否相似来判断是否是撒谎者，从而达到抵御攻击的目的。由于在系统中使用了时间折扣因子，即以前的实体的良好表现会随着时间的流逝而减少对现在判断的影响，所以该系统对叛国者攻击有一定的抵御能力。

4.2.3　Yan Lindsay Sun' model

Sun[23]等人提出了基于熵(entropy)理论的信任模型。该文主要比较了社会层面信任和计算机网络层面信任的差异，并总结了通过第三方建立信任(即信任繁衍)以及通过多源头建立信任(即多路繁衍)应该遵守的基本原理，为信任的建立制定了统一的规范，提出了在分布式环境下信任估算的基本框架。

该模型中采用 T(subject: agent, action)表示信任关系。Subject 是信任的主体(可以是一个实体，也可以是一个组织)；agent 是信任的客体(可以是一个实体，也可以是一个组织)；action 是客体的一个行为。并用 P(subject: agent, action)表示从主体的观点来看客体采取某行为的概率。基于熵的信任评估值定义如下：

$$T=\begin{cases}1-H(p), & 0.5\leqslant p\leqslant 1\\ H(p)-1, & 0\leqslant p\leqslant 0.5\end{cases} \tag{8}$$

其中 $p=P(subject: agent, action)$；$H(p)=-p\log_2(p)-(1-$

$p)\log_2(1-p)$

信任值的单路繁衍和多路繁衍具体含义如图 2 所示，其计算方法如下：

$$T_{ABC}=R_{AB}T_{BC} \tag{9}$$

$$T(A:C,action)=\omega_1(R_{AB}T_{BC})+\omega_2(R_{AD}T_{DC}) \tag{10}$$

其中 R_{AB} 为推荐信任，$\omega_1=\frac{R_{AB}}{R_{AB}+R_{AD}}$，$\omega_2=\frac{R_{AD}}{R_{AB}+R_{AD}}$。

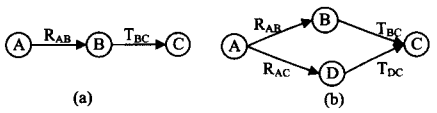

图 2　推荐信任融合示意图

为了体现客体行为随时间的变化情况，计算 $p=P(subject:agent,action)$ 时引入时间遗忘因子 $0\leqslant\beta\leqslant1$，其作用类似 Sonja Buchegger 模型中的时间折扣因子 u。具体计算公式如下：

$$P(A:X,action)=\frac{1+\sum\beta^{t_c-t_j}k_j}{2+\sum\beta^{t_c-t_j}N_j} \tag{11}$$

其中 t_c 为当前时刻，t_j 为统计时间，k_j 为行为执行的次数，N_j 为行为需要执行的次数。

该模型由于采用了信任与推荐信誉相分离的策略，所以对错误推荐攻击有较好的防御功能；其次，该文模拟部分详细分析了时间遗忘因子 β 的选取对叛国者攻击的检测的影响。如果 β 参数选择适当，该模型可以利用坏行为实体较难恢复良好信誉的方法，提高对叛国者攻击行为的抵御能力。该文对偏见式攻击做了一定文字分析，但并未阐述该模型对偏见式攻击具有定性或定量的良好表现。另外，该系统对联合式攻击依然没有较好的解决办法。

4.2.4　OTMF model

Li Ruidong[24] 等人针对偏见性攻击提出了一种基于贝叶斯方法的信任管理模型 OTMF，该模型融合了节点的自主信任观点和信誉信息，从结果可以看出对偏见性攻击有较好的抑制作用。

该模型使用 ITF 表示由原始数据产生的初始信任结构，例如 $ITF_{ij}=(\alpha_{ij},\beta_{ij})$，其中 α_{ij} 表示正确行为的次数，β_{ij} 表示错误行为的次数。为了抑制偏见性攻击，系统在计算信任评估时引入了信心值(confidence value)。信心值表示信任值计算的精确度，高信心值意味着目标实体已经经历了主体或者其他实体的多次检验。系统的执行过程可以分为以下 4 步。

S1：使用直接信息更新 ITF。更新方式如下：

$$\begin{aligned}\alpha_{ij}&=\alpha_{ij}+s\\ \beta_{ij}&=\beta_{ij}+1-s\end{aligned} \tag{12}$$

S2：分发和处理二手信息，各个节点周期性地根据直接信息产生二手信息并分发该信息。二手信息是从其他节点收集的关于目标节点的直接信息，更新方式类似 ITF 的更新方法。只是由于是经历了一个时间段，所以处理的是该时间段内正常行为和错误行为的次数。每一个节点都应该保留二手信息的来源以及二手信息的内容时间等信息。收到中间实体传来的二手信息要和本地的 ITF 信息比较，如果差异超过阀值就丢弃，该思想类似 Buchegger[7] 的做法。

S3：根据 S1 和 S2 步骤得到的 ITF 和二手信息，评估其他实体的基本观点由两个部分组成：信任值 $t\{i:j,action\}$ 和信心值(confidence value)$c\{i:j,action\}$，计算方法如下：

$$t\{i:j,action\}=E(Beta(x,\alpha,\beta))=\frac{\alpha}{\alpha+\beta} \tag{13}$$

$$\begin{aligned}c\{i:j,action\}&=1-\sqrt{12}\sigma(Beta(x,\alpha,\beta))\\&=1-\sqrt{\frac{12\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}}\end{aligned} \tag{14}$$

S4：评估可信度，根据信任和信心值计算可信度，计算方法如下：

$$T\{i:j,action\}=1-\frac{\sqrt{\frac{(t-1)^2}{x^2}+\frac{(c-1)^2}{y^2}}}{\sqrt{\frac{1}{x^2}+\frac{1}{y^2}}} \tag{15}$$

其中 x,y 是计算可信度的参数 。

该模型认为偏见攻击产生的主要原因是实体只考虑了自身的直接观察信息，而自身观察与实际目标实体的表现有偏差，所以导致偏见攻击。为了抵制这样的攻击，该系统考虑到将自身的观察和二手信息融合起来，并且在计算目标节点的可信度时，不仅考虑节点的信任值，还考虑衡量信任值的可信度的信心值。最后文章通过将 OTMF 模型与传统的两种方式建立的信任模型的比较，体现了该模型在抵御偏见攻击方面的优势。

由于 OTMF 借鉴了 Buchegger[7] 部分思想，所以该模型可以抵御错误推荐攻击，同时对偏见攻击做了相应的改进，所以对偏见攻击有一定防御能力。但该抵御能力是通过与最传统的两种模型做比较来体现的，并没有给出详细的定量分析。同样，OTMF 没有抵御联合攻击的能力。

4.3　各种模型的攻击防御能力及性能比较

本节主要讨论目前研究的自治网络中各种信任/信誉模型攻击防御所能达到的水平。具体各个系统设计思想以及对各种攻击防御能力在上两节做了一定的分析。为了能更好地体现各个系统在自治系统下使用的情况，表 1 除了比较各个模型在各种攻击抵御方面的能力，还比较了分布式计算需要的通讯量、本地信息存储资源、收敛性、可扩展性等自治网络环境下比较关注的性能的表现。

表 1　典型信任/信誉模型的攻击防御能力和性能比较

	Model	George	Tao Jiang	Buchegger	Sun	OTMF
攻击防御	错误推荐攻击	Better	Better	Better	Better	Better
	叛国者攻击	Lower	Lowest	Medium	Good	Medium
	偏见攻击	Lowest	Lowest	Lower	Medium	Good
	联合攻击	Low	Low	Lowest	Lowest	Lowest
性能	通讯量	Medium	Smaller	Medium	Medium	Larger
	收敛性	Lower	Good	Lower	Lower	Lower
	存储资源	Smaller	Smaller	Larger	Medium	Larger
	可扩展性	Lower	Good	Lower	Lower	Medium

George 模型是基于信任图路径搜索的思想，使用半环理论表达信任的计算模型。从模型的试验结果来看，对简单的错误推荐攻击具有较好的防御能力。作为流模型的一种，该模型和通常的流模型一样，没有考虑处理信任的时间特性，所以对叛国者攻击抵御能力弱。该模型讨论了极端式错误推荐，即坏节点采用对所有好节点推荐最小信任值、对所有坏节点推荐最大推荐值的策略，完全没有考虑偏见式攻击的防御问题。对于联合攻击，如果联合策略是单纯最大化周围坏节点信誉度、最小化周围好节点信誉度，则该系统具有一定的抵御能力。但对于复杂的联合攻击，该模型抵御效果有限。性

能方面，该模型是采用信任链的方法建立聚合推荐信任度，信任链跳数越多，计算收敛速度越慢；信任图的分布式存储依然是该模型的一个巨大挑战，从文献[15]看来，该模型采用集中存储，所以每个节点消耗的存储资源和通讯资源较少，但容易受到单点失效攻击。

Tao Jiang 模型采用了邻居节点直接投票方式建立信任评价，采用马尔科夫链分析各节点信任度的时间特征。通过模拟结果可以看出，该模型对错误推荐攻击和简单共谋攻击具有较好的防御能力，但对没有规律可循的偏见攻击和叛国攻击，信任评价正确性将随攻击者增加而急剧下降。在性能方面，该系统采用了邻居节点直接投票方式，所需保存信息仅为周围节点信息，且只需与周围节点通讯，所以通讯量和对存储资源的需求将较小；其次，由于采用了分布式迭代算法，具有较好的可扩展性，并且作者在文献中证明了该模型具有较好的收敛性。当然，该模型存在信任度量以及数学分析使用的网络环境过于简单化的问题。

Buchegger 模型采用了贝叶斯理论，通过获得的第一手资料和自身保存信息比较识别撒谎者。从模拟结果可以看出，对简单的错误推荐攻击具有较好的抵御功能。但是在撒谎者较多的环境下，抵御能力有所下降。由于系统使用了时间折扣因子，所以对叛国者攻击具有简单的防御能力；对于偏见攻击，文章没有做任何讨论。根据该系统的工作原理来看，该系统对偏见攻击难以识别；当多个撒谎者作出协同的攻击，该系统难以抵御。性能方面，由于该系统需要在每个节点保存它关于其他节点 R_{ij}、T_{ij} 和 F_{ij}，所以所需的存储资源较大；由于节点之间只需交互 F_{ij}，但几乎所有节点都有交互 F_{ij} 信息的可能，所以通讯量中等；新节点加入时，尚不能和其他节点交互 F_{ij} 信息，即建立信誉评价不准确，并需要较长时间建立正确的信誉评价，所以可扩展性与计算收敛性也是一个问题。

Sun 模型提出了基于熵(entropy)理论的信任模型，专门针对信誉系统的欺骗问题，采用了信任与推荐信誉相分离的策略，所以对错误推荐攻击有较好的防御功能；模型中采用了时间遗忘因子，并通过对遗忘因子的调整可以达到抵御叛国者攻击的效果；文章中对偏见式攻击做了文字性的分析，但并未看到定性或定量的效果；对于联合式攻击，文章没有提及。从其工作的基本原理来看，对联合式攻击抵御能力不会强于文中的其他模型。性能方面，由于每个节点仅保存有关其他节点的信誉度和信任度，所需存储资源比 Buchegger 模型少；节点之间只需要传递推荐信息，所以通讯量和 Buchegger 模型类似；由于该模型采用的信任建立方法和 Buchegger 模型类似，因此扩展性和收敛性依然是一个问题。

OTMF 模型是针对偏见式攻击提出的一种基于贝叶斯方法的信任管理模型，该模型融合了节点的自主信任观点和信心信息。试验结果表明，该模型对错误推荐攻击具有一定的防御能力；由于结合了自主观点与信誉信息，该模型获得的信任评价更加客观，所以对偏见式攻击也具有较好的防御能力；由于该模型借鉴了 Buchegger 模型思想，采用了时间折扣因子，所以对叛国者攻击具有一定的抵御能力；同样，该模型未进行联合攻击方面的研究，对于联合攻击的抵御依然薄弱。性能方面：各节点所需保存的信息量与 Sun 模型类似，但 OTMF 模型在传递第一手信息的同时还需要周期性地分发第二手信息，所以通讯量较 Sun 模型大；计算信任评价需要融合信任度和信心信息，由于信心信息是全局评价，所以该模型的可扩展性和计算收敛性依然没有得到较好解决。

5 存在的问题与展望

5.1 当前研究存在的问题

关于自治网络中的信任/信誉模型本身的安全问题已经受到一些关注，但是通过上节的分析，可见还存在一些明显的问题。

• 信任度量定义的混乱：目前信任关系依然是社会的最复杂的关系之一，这也是一种主观化的认知。虽然 Sun[23] 提出了计算机网络中的信任与社会信任不同，并分析了它们之间的差异，但并没有得到广泛的认同。因此信任度量的方法繁多，各种不同的模型都根据自己的需要定义不同的度量标准，这给以后评估工作带来很大困难。

• 攻击防御的局限性：从表 1 可以看出，目前大多数模型对最简单的错误推荐攻击有防御能力，对其它攻击基本不能同时做到较好的防护，对联合攻击的研究没有涉及。

• 攻击防御能力无统一评价标准：目前大多数模型在评价自己对某些攻击的防御能力时，所做分析大多是在自身预先定义的基础之上的结果展示，很少与其他模型进行效果对比。主要是因为各个模型对攻击的定义、防御目的和效果各自有不同的定义，难以做同类型比较，甚至比较也是无意义的。

• 信任模型性能评价困难：各种模型都是基于不同的应用背景提出来的。虽然自治网络具有它自身的特点，但是各个模型在满足自治网络的不同需要方面使用不同的手段，达到不同的平衡。虽然现有信任/信誉模型都声称可以适用于自治网络，但是目前没有任何信任模型做过相关性能分析。

5.2 展望

结合第 5.1 节中提出的问题，我们认为在自治网络中信任/信誉模型的安全问题，可以在以下几个方面做一些研究工作：

• 进一步研究信任关系。究竟建立什么样的信任关系以及采用何种信任度量方法可以较好地适应自治网络中的普遍使用需求。

• 对于所有攻击的防御问题，更是研究信任/信誉模型安全问题的重点。信任关系和信任模型的性能评价是为安全问题打下基石，对所有攻击的全面防御是以后研究工作的主要内容。

• 探索解决安全问题的新途径。本文所讨论的基本是通过安全模型本身的改进来提高模型对攻击的抵御能力，结合其他学科的知识探索新的途径也许能达到更好的效果。

• 信任模型的性能评价问题。如何对众多自治网络中信任模型进行客观的性能评价分析，是需要解决的一个重要问题。

参 考 文 献

[1] IETF Mobile Ad-hoc Networks (manet) working group. [Online]. http://www.ietf.org/html.charters/manet-charter.html

[2] Gnutella. [Online]. http://www.gnutella.com

[3] Stoica I, Morris R, Karger D, et al. Chord: A scalable peer-to-

peer lookup service for internet applications// Proceedings of the ACM SIGCOMM '01 Conference. San Diego, California, August 2001:149-160

[4] Gupta R, Somani A K. Reputation Management Framework and Its Use as Currency in Large-scale Peer-to-Peer Networks // Fourth International Conference on Peer-to-Peer Computing (P2P'04). 2004

[5] Dewan P, Dasgupta P. Securing P2P networks using peer reputations: Is there a silver bullet // IEEE Consumer Communications and Networking Conference(CCNC). 2005

[6] Marti S, Garcia-Molina H. Limited Reputation Sharing in P2P Systems// ACM Conference on Electronic Commerce (EC'04). 2004

[7] Buchegger S, Le Boudec Jean-Yves. A Robust Reputation System for P2P and Mobile Ad-hoc Networks. P2PEcon, 2004

[8] Jaramillo J J , Srikant R . DARWIN : Distributed and Adaptive Reputation mechanism for WIreless ad-hoc Networks. MobiCom'07. Montréal, Québec, Canada, September 2007

[9] Olmedilla D, Rana O, Matthews B, et al. Security and trust issues in semantic grids // Proceedings of the Dagsthul Seminar. Semantic Grid: The Convergence of Technologies, 2005, 05271

[10] Aringhieri R, Damiani E, Paraboschi S, et al. Fuzzy Techniques for Trust and Reputation Management in Anonymous Peer-to-Peer Systems. Journal of the American Society for Information Science and Technology, 2006, 57(4)

[11] Lamsal. Understandingtrustandsecurity. http://www. cs. helsinki. FI/u/lamsal/papers/

[12] Romano D M. The Nature of Trust : Conceptual and Operational Clarification. PhD Thesis. Louisiana State University, 2003

[13] JØsang A. Trust and Reputation Systems. Foundations of Security Analysis and Design IV, FOSAD 2006/2007 Tutorial Lectures. Springer LNCS 4677. Bertinoro, Italy, September 2007

[14] JØsang A, Ismail. A survey of trust and reputation system for online service provision. Decision Support Systems, 2007, 43(2): 618-644

[15] Theodorakpoulos G, Baras JS. On Trust models and trust evaluation metrics for ad-hoc networks. IEEE Journal on Selected Areas in Communications, 2006, 24(2): 318-328

[16] Theodorakpoulos G. Distributed trust evaluation in ad-hoc networks. MS Thesis. 2004

[17] Page L, Brin S, Motwani R, et al. The PageRank Citation Ranking: Bringing Order to the Web. Technical report. Stanford Digital Library Technologies Project, 1998

[18] Ziegler C-N, Lausen G. Spreading Activation Models for Trust Propagation// Proceedings of the IEEE International Conference on e-Technology, e-Commerce, and e-Service (EEE '04). Taipei, March 2004

[19] Levien R. Attack Resistant Trust Metrics. PhD thesis. University of California at Berkeley, 2004

[20] Newsome J. The Sysbil Attack in Sensor Networks: Analysis & Defenses// the Third International Symposium on Information Processing in Sensor Networks (IPSN). April 2004

[21] Wang YuFeng , Hori Yoshiaki . On securing open networks th - rough trust and reputation-architecture, challenges and solutions. Technical Report. Institute of Electronics, Information and Communication Engineers, 2006, 106(340): 1-6

[22] Danezis G, Schi_ner S. On network formation, (Sybil attacks and Reputation systems). http://dimacs. rutgers. edu/Workshops/InformationSecurity/slides/gamesandreputation. pdf

[23] Sun Y, Yu W, Han Z. Information theoretic framework of trust modeling and evaluation for ad hoc networks. IEEE Journal on Selected Areas in Communications, 2006, 249(2): 674-679

[24] Li Ruidong, Li Jie, Liu Peng, et al. An Objective Trust Management Framework for Mobile Ad Hoc Networks // Vehicular Technology Conference, 2007. 2007: 56-60

[25] Jiang Tao, Baras J S. Trust Evaluation in Anarchy: A Case Study on Autonomous Networks // Infocom. Barcelona, Spain, April 2006

[26] Mundinger J, Jean-Yves. Reputation in Self-organized Communication Systems and Beyond// Proceedings from the 2006 Workshop on Interdisciplinary Systems Approach in Performance Evaluation and Design of Computer & Communications Systems. Pisa, Italy, 2006

[27] Mundinger J , Jean - Yves . Analysis of a Reputation System for Mobile Ad-hoc Networks with Liars// Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks, 2005, Third International Symposium on. 2005: 41-46

[28] Mundinger J, Jean-Yves. The Impact of Liars on Reputation in Social Networks. Social Network Analysis: Advances and Empirical Applications Forum, 2005

(上接第 4 页)

[22] Cai Jun, Tan K, Ooi B. On Incremental Cache Coherency Schemes in Mobile Computing Environment// Proc. of 13th Intl. Conf. Data Engineering. 1997: 114-123

[23] Wu K, Yu P, Chen M. Energy-Efficient Caching for Wireless Mobile Computing// Proc. of 20th Intl. Conf. Data Engineering. 1996: 336-343

[24] Tan K, Cai Jun, Ooi B. An Evaluation of Cache Invalidation Strategies in Wireless Environments. IEEE Trans. Parallel and Distributed Systems, Aug. 2001, 12(8): 789-807

[25] Lan Jiang, Liu Xiaotao, Shenoy P, et al. Consistency Maintenance in Peer-to-Peer File Sharing Networks// Proc. of the 3rd IEEE Workshop on Internet Applications. 2003

[26] Tan K, Cai J. Broadcast-based Group Invalidation: An Energy Efficient Cache Invalidation Scheme. Information Sciences, 1997, 100(1): 229-254

[27] Huang Yu , Jin Beihong , Cao Jiannong , et al . A Selective Push Algorithm for Cooperative Cache Consistency Maintenance in MANETs// IFIP Intl. Conf. on Embedded and Ubiquitous Computing (EUC), 2007

[28] Sailhan F, Issarny V. Cooperative Caching in Ad hoc Networks // IEEE Intl. Conf. on Mobile Data Management (MDM). 2003

[29] Cao Jiannong, Zhang Yang, Xie Li, et al. Consistency of Cooperative Caching in Mobile Peer-to-Peer Systems Over MANET// IEEE Intl. Conf. on Distributed Computing Systems Workshops. 2005

[30] Huang Yu, Cao Jiannong, Jin Beihong. A Predictive Approach to Achieving Consistency in Cooperative Caching in MANET // Proc. of the 1st Intl. Conf. on Scalable Information Systems, P2PIM workshop session, ACM Intl. Conf. 2006