Journal of Computer Applications
计算机应用,2016,36(10):2715 -2722
ISSN 1001-9081
CODEN JYIIDU
2016-10-10
http://www.joca.cn

文章编号:1001-9081(2016)10-2715-08
DOI:10.11772/j.issn.1001-9081.2016.10.2715

面向云数据安全自毁的分布式哈希表网络节点信任评估机制

王　栋[1,2],熊金波[1,3*],张晓颖[1]

(1.福建师范大学 软件学院,福州 350117;　2.厦门大学 软件学院,福建 厦门 361005;
3.福建省公共服务大数据挖掘与应用工程技术研究中心,福州 350117)

(*通信作者电子邮箱 jinbo810@163.com)

摘　要: 在云环境下的数据自毁机制中,针对分布式哈希表(DHT)网络的恶意节点和不诚信节点容易造成密钥分量丢失或泄露等问题,提出面向云数据安全自毁的DHT网络节点信任评估机制。该机制首先为节点建立二维信任评估模型,对节点可信程度进行定性和定量的考察;接着改进节点直接信任值和推荐信任值的计算方法,充分考虑节点内外因素,从节点运行环境和交互经验两个维度出发并细化到不同层级计算节点信任值,得到信任评价子云;然后将各个信任评价子云加权合成得到综合信任云,采用云发生器描绘出综合信任云一维高斯云图形;最后结合信任决策算法选出最优可信节点。实验结果表明该机制能够帮助原有的数据自毁机制找到更适合存储密钥分量的DHT网络节点,提高了系统的容灾能力并降低了系统计算负载。

关键词: 云数据;数据自毁;分布式哈希表;信任评估;信任云

中图分类号: TP309.2　　**文献标志码:** A

Trust evaluation mechanism for distributed Hash table network nodes in cloud data secure self-destruction system

WANG Dong[1,2], XIONG Jinbo[1,3*], ZHANG Xiaoying[1]

(1. *Faculty of Software, Fujian Normal University, Fuzhou Fujian* 350117, *China*;
2. *Faculty of Software, Xiamen University, Xiamen Fujian* 361005, *China*;
3. *Fujian Engineering Research Center of Public Service Big Data Mining and Application, Fuzhou Fujian* 350117, *China*)

Abstract: Distributed Hash Table (DHT) network is widely used in secure self-destruction schemes of privacy data in cloud computing environment, but malicious nodes and dishonest nodes in the DHT network easily lead to key shares loss or leakage. To tackle those problems, a trust evaluation mechanism was proposed for the DHT network used in cloud-data secure self-destruction system. In this mechanism, a trust cloud model was established for DHT nodes to describe their trust information qualitatively and quantitatively. By introducing an improved calculation method of direct trust value together with recommended trust value and fully considering the internal and external factors of DHT nodes, the trust value of nodes were first calculated on two dimensions consisted of operating experiment and interactive experience. The result data were used to build trust evaluation sub-cloud for each index. After that, all these trust evaluation sub-clouds were summed up to generate the comprehensive trust cloud according to the weights of different evaluation indexes. Then, the comprehensive trust cloud, by means of cloud generator algorithm, could be described as one-dimensional normal cloud. Finally, the reliable and efficient nodes could be selected using trust decision algorithm. Experimental results show that the proposed mechanism can help original data self-destruction system making comprehensive trust decision and finding reliable DHT network nodes, further enhancing disaster-tolerant capability and reducing computational cost of the system.

Key words: cloud data; data self-destruction; Distributed Hash Table (DHT); trust evaluation; trust cloud

0　引言

随着云时代的来临和数据科技的不断发展,人们创造、分享和存储数据的热情空前高涨。云存储技术的出现极大地满足了人们对海量数据存储的需求,各式各样国内外云存储产品也逐渐被人们所依赖,如EMC Atmos、IBM XIV、iCloud、SkyDrive、Dropbox、百度云、华为网盘等。然而,在开放互联网环境中,这些存储在云端的数据由于其管理权与所有权分离,更容易受到黑客攻击或人为泄露,数据的泄露和破坏等安全问题必将对企业或个人造成重大的损失[1]。

数据安全性问题制约了云存储技术的发展,最直接的影响是产品失去黏性,造成一批僵尸用户或准僵尸用户的产生。据统计,大约有80%的企业不愿将自己的业务数据、机密文件存储在云存储产品,普通用户一般也只愿存放视频、图片、音乐等重要性程度较低的内容。因此,在新业务场景需求的推动下,云环境下的数据存储需要一种面向重要数据保护的

收稿日期:2016-04-08;修回日期:2016-06-15。　基金项目:国家自然科学基金资助项目(61402109,61370078);福建省自然科学基金资助项目(2015J05120);福建省高校杰出青年科研人才培育计划项目(2015)。

作者简介:王栋(1993—),男,福建福州人,硕士研究生,主要研究方向:信任评估、数据安全;　熊金波(1981—),男,湖南益阳人,副教授,博士,CCF会员,主要研究方向:云数据安全、隐私保护;　张晓颖(1995—),女,福建泉州人,主要研究方向:信任评估、数据安全。

安全销毁机制。这种机制的存在使得高质量的云存储服务得以实现,同时可解决动态共享和安全性的矛盾问题。

云环境下数据安全销毁早期研究方向按密钥管理体制不同可分为集中式解决方案和分散式解决方案[2]。集中式解决方案依赖于可信第三方,典型例子有:Perlman等[3]设计了一种系统,该系统以删除过期加密密钥的方式使得加密文件不可恢复。在此基础上,FADE(File Assured DEletion)系统[4]采用基于策略的访问控制并对数据进行二次加密实现数据安全销毁;Ephemerizer系统[5]则可以允许数据拥有者自主设计密钥的过期时间。但可信第三方的引入,大大增加了集中式数据销毁方案的潜在风险。为避免这种人为因素的影响,分散式解决方案不引入任何可信第三方,而是充分利用分布式网络的特性实现数据自毁,此类型的典型代表有:Geambasu等[6]提出的Vanish系统利用门限秘密共享和分布式哈希表(Distributed Hash Table,DHT)网络节点周期性更新的特点实现密钥分量的自动销毁。文献[2,7-9]中设计的系统分别采用不同的公钥加密机制对Vanish系统作出改进,然而,DHT网络具有不稳定性,本身容易遭受跳跃攻击和嗅探攻击等Sybil攻击,均被证实存在一定的安全隐患[10]。

熊金波等[11]在分析总结了国内外数据销毁方案的局限性后,结合多级安全、DHT网络和基于身份的加密(Identity-Based Encryption,IBE)提出基于身份加密的安全自毁方案。该方案先将隐私内容分为不同安全等级模块并进行加密得到密文,再对所得密文进行IBE和处理得到混合密钥分量,然后将混合密钥分量散发到DHT网络中。方案改善了复杂密钥管理和分发的问题,同时也提出了细粒度访问控制策略,但没有考虑DHT网络节点的可信度问题。为此,吴守才等[12]提出基于信任评估的数据自销毁方案,其核心思想是首先使用密钥派生树加密隐私数据,然后为DHT网络节点建立信任模型,最后将密钥分量分发到信任值高的节点,依赖节点的周期性更新实现自销毁。该方案通过考察节点的可信度,提高节点服务质量进而保证了方案的安全性。但该方案亦存在以下几个问题:1)信任评估指标不够全面。该方案只从节点的交互经验层面上进行计算,忽略了对节点内部属性即节点性能、流量参数、生存时间参数的考虑,评价指标比较单一;2)节点的信任值是经过对评估指标迭代计算后存储在DHT网络中,因此随着DHT网络节点的周期性更新,节点的历史信任值数据也随之删除。

从Vanish系统提出开始,DHT网络便广泛应用在各种数据自毁机制中,可见其已成为数据自毁机制中不可或缺的一部分。DHT网络周期性自更新特性是一个优势,但同时也存在不稳定性,这意味着混合密钥分量一旦被分发到恶意节点或不可信节点,将会导致它们还在生存期内就失去价值,还可能成为攻击对象。此类节点一旦增多,系统将承担更大的风险。因此,在数据自毁机制中,为提高整体服务质量,给DHT网络节点建立信任评估机制十分必要。针对上述问题,本文在总结前人研究的基础上,结合信任云[13]建立信任模型,提出面向云数据安全自毁的DHT网络节点信任评估机制。本文的主要工作如下:

1)针对DHT网络节点,实现二维信任云评估模型的构建,不仅能够定性和定量的描述节点间的“信任”关系,还可以迭代计算并存储信任值,提高信任值的准确度。

2)提供更加全面的评估指标,综合节点的内部属性(节点的运行环境)和外部表现(节点间的交互经验)进行节点的可信度计算。

3)为系统信任决策提供依据,帮助系统选择可信度更高的节点进行存储混合密钥分量,降低由DHT的不稳定性带来的潜在风险。

1 相关知识

1.1 DHT网络

DHT网络通常被认为是分布式系统[14],其本质上是一种数据结构,通过分布式哈希函数实现数据的分布式存储,是P2P网络中用于资源组织和查找的底层架构。DHT网络的实现方法多样因此其种类也很繁多,常见的如Vuze、Kademlia(简称KAD)和OpenDHT等。所有的DHT网络都有以下三个特性[11]:

1)DHT网络具有可用性。DHT网络可用性不仅体现在分布式存储机制使得同一数据可以同时存储在不同的节点上,还体现在其支持节点动态进入/退出、自组织能力以及良好的扩展性。

2)DHT网络节点周期性更新。DHT网络为用户提供一个稳定的节点更新周期,一旦节点上的数据超过存储期限便会自动删除,继而存储新的数据。不同的DHT网络更新周期各不相同,如Vuze DHT的更新周期是8 h,OpenDHT可在一周的范围内根据用户需要自主设定。

3)DHT网络节点规模大且分布全球化。文献[15]指出,在Vuze DHT网络中活跃节点规模达到百万以上且属于全球不同的国家和地域。

在云数据安全自毁方案中,DHT网络节点通常用于存储用户隐私内容的密钥分量,其第一个特性是适用于构建数据自毁方案的基本条件;第二个特性满足用户在没有任何可信第三方或人为干预下数据自动删除的要求;第三个特性保证DHT网络节点能够抵抗各种针对性的非法攻击。

1.2 信任及其属性

信任是一个复杂的主观概念,其本身含义的不确定性使得不同学科中信任的含义表现也各不相同[16]。在社会科学领域中,信任是一种有价值的人际关系;信息技术领域中,对信任评估机制的仿照真实社会中人们获得声誉的方式以及判断的过程;在DHT网络中,信任是指对节点一次或多次交互行为的肯定,意味着节点能使得混合密钥分量在其生存期内保持价值。信任有以下几个属性[13,17-18]:

1)主观性,信任的评估容易带有评价主体的个人色彩,无法确保客观;

2)多相关性,信任与多种属性相关,如所处环境、交互情况等;

3)动态性,信任评估结果因时间、环境等外因的变动而变动;

4)有限传递性,信任值在传递过程中,随着中间节点的增多而递减;

5)可度量性,通过模糊理论和概率论知识可对信任进行度量;

6)有向性,节点P对节点Q的评估是单向的,评估结果不能反向适用。

1.3　云的基本概念

设 U 是一个用数值集合表示的论域 $U=\{x\}$，$\tilde{A}$ 是 U 上的某一定性概念，x 为论域 U 中的元素且 x 是定性概念 $\tilde{A}$ 在数量上的一次随机实现。若任取一个 x 的值都有一个稳定倾向的随机数 $y=\mu_{\tilde{A}}(x)\in[0,1]$ 与其对应，则称 x 在论域上的分布范围为隶属云，简称云。每一对确定的 (x,y) 值是云的一个云滴，大量服从稳定分布的云滴构成了云。云的形态由它的三个特征参数 (Ex,En,He) $(0<Ex<1,0<En<1,0<He<1)$ 反映。其中，Ex 表示基本信任值; En 是信任熵，表示信任的不确定性; He 反映熵的不确定性。

本文将信任这个不确定性概念以高斯云(正态云)形式表达出来，借助信任云描述 DHT 节点间的信任关系。

1.4　信任云模型

信任云模型[19]是基于云理论[20]提出的一种特殊的云模型。该模型通过定义一个包含五个因素的信任向量为分布式节点处理信任数据，再借助云发生器将所有数据记录进行云化，刻画出云图形; 最后根据云图形所描述的信任关系，借助预设的基准云和信任决策算法作出信任决策。文献[13,18,21]在此思路上提出了改进后的信任云模型。

本文中引入信任云模型对 DHT 网络节点进行可信性评估，原因如下:

1) 以正态云的形式代替精确值表示信任的概念，能够更全面地把握信任概念中存在的随机性和模糊性，有助于更加科学地进行信任决策。

2) 结合云计算平台为节点提供高效的信任值计算和存储机制。DHT 网络节点的信任值不再保存在本地: 一方面，强大的云计算能力有助于减少对信任值不断迭代计算带来的时间开销; 另一方面，节点的信任值得以长期保存。

3) 拓展性良好。信任评估模型支持从多个维度全面的计算节点的信任值，也可根据策略需要添加、删除、更新信任子云。

2　二维信任云评估机制

在数据自毁机制中，为 DHT 节点建立合理的信任评估机制十分重要。由于目前网络环境的复杂性，对节点的评估应建立在综合考虑其内外部因素的基础上，也就是说，对节点信任值的考察，不仅需要考虑 DHT 节点交互满意度情况，更要结合节点的网络环境作出综合评估。为实现上述系统要求，本文结合信任云，将所提的面向云数据自毁的二维信任云评估机制描述如下。

2.1　相关定义

定义 1　直接交互信任度 DTrust。指 DHT 网络节点间产生直接交互行为并对其进行评价得到的信任度，是客观的评价节点是否成功地完成某项交互动作的可信程度。

定义 2　推荐交互信任度 RTrust。指推荐节点 P 基于历史交互行为或主观感受，间接为第三方提供节点 Q 的推荐值，可有限传递且具有多相关性。

定义 3　交互经验信任度 Trust。DHT 网络二维信任评估体系的评估维度之一，是对节点外部工作表现的考察，表示节点全局信任度值，由上述直接交互信任度和推荐交互信任度按照一定的权重比得到的综合评价，可预测该节点未来的信任状态。

定义 4　运行环境信任度。DHT 网络二维信任评估体系的评估维度之一，是对节点所在的 DHT 网络环境进行评估而得到的对该节点性能(如节点生存参数、流量参数、搜索性能等)及安全方面的可信度预测(自身性能表现)。

2.2　核心思想

为得到可信 DHT 网络服务节点建立信任云模型，当外部数据对节点有交互意向发生时，该模型从节点的不同评估维度计算和收集信任值并云化，建立信任评价子云; 根据各维度重要级别不同，依次进行合成得到综合信任云; 最后，结合云图形依照信任决策算法选择符合信任要求的节点。模型的评估过程见图 1，具体描述如下:

1) 划分信任等级，针对各个信任级别设定标度范围。

2) 依照评估指标的设定，采集有交互经验的节点的有关数据(运行环境信任度和交易经验信任度，包括性能表现、安全表现、直接推荐信任度和推荐信任度)。然后将得到的数据进行初始化。本文采用离散标度的形式处理数据: $V(a_1,a_2,a_3,a_4,a_5)$，V 为信任向量。

3) 将各个维度的信任向量经逆向云生成器转换为信任云参数。通过正向云生成器，用信任云参数描画出信任云图形。

4) 将各个维度的信任评价云按照权重系数组合，得到综合信任云。

5) 将得到的综合信任云观察云形状并与预先设定的基准云进行对比，最后通过信任决策算法选出可信节点。

图 1　二维信任云评估机制工作过程

2.3　云发生器

云发生器是正向云发生器和逆向云发生器的统称，是云模型的重要组成部分。在信任云模型中，逆向云发生器算法将信任向量生成三个云特征参数，见算法 1; 正向云生成器算法利用云的三个特征参数作为输入值描绘出符合正态分布的云图形，见算法 2。

算法 1　逆向云生成器。

输入　样本数据 $x_i, i\in[1,N]$。

输出　信任云的特征参数 Ex、En、He。

for $i = 1: N$ // 计算样本数据平均值、标准差、样本方差

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i, std = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2}$$

$$S^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2$$

End

$Ex = \bar{x}$;

$En = std$;

$He = (S^2 - En^2)^{\frac{1}{2}}$

算法 2 正向云生成器。

输入 特征参数 Ex、En、He，以及常数 N。

输出 N 个云滴。

for $i = 1: N$ // 产生 N 个云滴

$Enn = \text{normrnd}(En, He)$;

// 产生一个 En 为期望值，He^2 为方差的正态随机数 Enn;

$Enn = \text{abs}(Enn)$;

$x_i = \text{normrnd}(Ex, Enn)$;

// 产生一个 Ex 为期望值，Enn 为方差的正态随机数 x_i

if $Enn \sim = 0$ // 计算确定度 μ_i

$\mu_i = \exp(-(x_i - Ex)^2/(2 * Enn^2))$;

else

$\mu_i = 1$;

end

end

2.4 信任参数的获取

节点的可信度由节点间的交互结果决定。节点的信任信息用数值进行量化得到信任值，取值范围为 [0,1]。按照信任程度的高低将其划分成五个信任级别，用自然语言表述为十分可信、可信、一般、不可信、十分不可信，不同的信任级别对应不同的信任区间。本文采用离散标度的方式定义一个由五个元素组成的信任向量 $V(a_1, a_2, a_3, a_4, a_5)$，每个元素分别表示对五个信任级别在对应信任区间内的统计次数。经过多次实验并统计分析，本文在使用信任向量计算时用 (1, 0.75, 0.5, 0.25, 0) 分别对应上述五个信任级别，如表 1 所示。

表 1 信任级别描述及其标度

信任级别	可信度描述	取值
1	十分不可信	0.00
2	不可信	0.25
3	一般	0.50
4	可信	0.75
5	十分可信	1.00

在信任评价过程中，专家对某个节点的评估信息可用信任向量表示。例如：专家 p 对节点 q 的 100 次服务行为评分，其中 10 次评为十分可信，20 次评为可信，70 次评为一般，则 p 对 q 的评估结果可用信任向量表示为 $V(10, 20, 70, 0, 0)$。

2.5 运行环境信任度计算方法

为了更好地评估目标节点的可信程度，需针对 DHT 网络节点制定更全面的评估指标，对其特征行为进一步细化。而在 DHT 网络中，节点的性能表现能够影响其服务质量。因此，本文在考虑节点交易经验信任度的同时，引入对节点网络运行环境的考察，使信任云模型的评价体系能够全面反映节点的特性。根据 DHT 网络的特点，本文构建了节点运行环境信任评价指标体系，如图 2 所示。

在运行环境维度的指标体系中，各个指标对受评价对象影响程度不同，对应的权重也各不相同。记权重系数为 ω，$\omega = \{\omega_i > 0, \sum_{i=1}^{n}\omega_i = 1\}$。权重系数可由评价者自行设定，也可以根据模糊层次分析 (Fuzzy Analytic Hierarchy Process, FAHP) 法[23] 确定。

设节点生存参数评价云为 $TL(Ex, En, He)$，流量参数评价云为 $TF(Ex, En, He)$，搜索性能评价云为 $TI(Ex, En, He)$，节点物理能力评价云为 $TP(Ex, En, He)$，安全性评价云为 $TS(Ex, En, He)$，综合信任评价云为 $T(Ex, En, He)$，则运行环境信任度可由式 (1) 计算：

$$T = \omega_1 \times TL + \omega_2 \times TF + \omega_3 \times TI + \omega_4 \times TP + \omega_5 \times TS \quad (1)$$

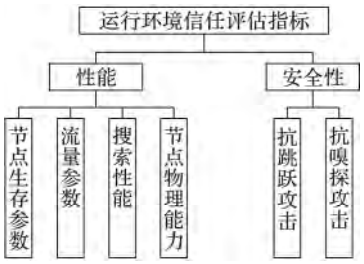


图 2 运行环境评估指标

2.6 交互经验信任度计算方法

交互经验信任度由每一次交互的评价累加得到。二维信任云模型中，交互经验信任度值由两个部分计算结果组成：1) 计算节点自身的直接交互经验信任值；2) 计算该节点的推荐节点对其的推荐值。综合交互经验信任度可由式 (2) 计算：

$$Trust = \alpha DTrust + (1 - \alpha) RTrust;\ \alpha \geqslant 0 \quad (2)$$

其中：$DTrust$ 为直接交互经验信任值；$RTrust$ 为间接信任值；α 为平衡因子，用户可按照不同的信任策略设定。

2.6.1 直接信任值计算方法

计算节点的直接信任度值需要考虑：惩罚前综合时间衰减、交易频量因素影响的信任评价值以及与服务失败次数相关的惩罚值。直接信任值的计算可用公式表示为：

$$DTrust = Trustsub - Punish(i) \quad (3)$$

其中：$Trustsub$ 表示惩罚前结合时间衰减因素、交易频量因素计算的信任值；$Punish(i)$ 为节点服务失败的惩罚值。

定义 5 时间衰减因子 ft。考虑信任值的时效性，通过时间衰减函数实现信任值受时间属性影响逐渐降低。信任值的有效性与时间跨度负相关。当一个节点的受评价时刻与当前时刻相距越近越能反映该节点服务行为的可信程度，信任值的有效性越高；相反，有效程度越低。ft 的计算公式如下：

$$ft = e^{-k(t_i - t_0)} \quad (4)$$

其中：t_i 指当前时刻；t_0 指节点的建立时刻；k 为平衡系数。

定义 6 交易频量因子 λ。节点的服务次数能够反映节点的受信赖程度。信任值与节点的交易频量正相关，节点服务频率越高，信任值越大，该节点越可信。λ 值计算公式如下：

$$\lambda = \frac{n}{n + C} \quad (5)$$

其中：n 表示某节点的服务次数；C 为可人工设置的常数，本文取多次实验后的最优值 0.4。

局部信任值的计算公式可表示为：

$$Trustsub = \sum_{i=1}^{I(u)} \frac{S(i) \times \lambda \times ft}{I(u)} \quad (6)$$

其中: $S(i)$ 为评价者对节点的评分，取值为(0,0.25,0.5,0.75,1); $I(u)$ 为节点的服务次数。

定义 7　惩罚因子。当节点的服务失败时，以减少信任值的方式对其进行惩罚处理。惩罚因子根据节点服务失败的频次，可对信任值进行不同程度的惩罚。节点失败频次越多，惩罚值越大。记惩罚因子为 θ，则:

$$\theta = \kappa \times \frac{1}{1 + \ell^{-n}}$$

其中: k 为人工设置的常数; n 为失败次数;

$$k \in \begin{cases} (0,0.3], & n \leqslant 5 \\ (0.3,1], & n > 5 \end{cases}$$

$$Punish(i) = \frac{1}{I(u)} \sum_{i=1}^{I(u)} S(i) \times \theta \tag{7}$$

综上所述，节点的直接信任度可由式(8)计算:

$$DTrust = \sum_{i=1}^{I(u)} \frac{S(i) \times \lambda \times ft}{I(u)} - \frac{1}{I(u)} \sum_{i=1}^{I(u)} S(i) \times \theta \tag{8}$$

2.6.2　三间接信任值计算方法

节点 p 的间接信任度主要计算思路为同时以推荐者的信任值和 p 所在 DHT 网络的邻居节点对 p 的推荐评价为重，强调推荐可靠性。本文选取的推荐节点是指与受评节点 p 有交互历史的节点，不考虑通过多条路径推荐方式。计算间接信任值的依据主要有以下几个方面:

1) 节点 p 的推荐节点对 p 的推荐值。推荐节点对 p 给出的评估值越大，节点 p 越可信。

2) 节点 p 的推荐节点数目。推荐节点的数量越多，越有利于提高间接信任值计算的准确性。

3) 推荐节点自身的直接信任度值。它是评估推荐可靠性的主要参考依据，推荐节点自身的信任度值越高，推荐值越可靠。

4) 推荐节点对节点 p 的推荐时间跨度。引入时间衰减因子，推荐节点对 p 的推荐时刻与当前时刻跨度越小，推荐值的有效性越大。

定义 8　推荐信任权重。该权重与推荐节点的推荐值及时间衰减因子有关。权重的值越大，表示节点越可信。反之，节点越不可信。记推荐信任权重为 Rv_i，计算公式如下:

$$Rv_i = \ell^{-1} \Big/ \sum_{k=1}^{I(k)} S(k) \times f_t \tag{9}$$

其中: $S(k)$ 为历次推荐节点给出的推荐值; f_t 表示时间衰减因子。

间接信任值由全部推荐信息经计算后取平均得到，将其计算公式定义如下:

$$RTrust = \frac{1}{N_{\text{recom}}} \sum_{i=1}^{N_{\text{recom}}} S(i) \times Rv_i \tag{10}$$

其中: N_{recom} 为节点 p 推荐节点的总数目; $S(i)$ 为推荐节点自身的直接信任度值; Rv_i 为推荐信任权重。

2.7　信任评价子云的合成

在二维信任评估体系中，每个维度往往包含一个或多个子维度即信任评价子云，云的合成指的是将两个或是两个以上的信任评价子云按照特定的权重系数合成，得到新的信任评价云; 从最底层的细分维度开始，然后依次向上一层维度合成最终形成全局信任综合评价云。权重系数按照各个评估维度的重要程度不同进行区别设定。记权重系数为 λ，$\lambda = \{\lambda_i > 0, \sum_{i=1}^{n} \lambda_i = 1\}$。$\lambda$ 可由评价者自行决定，也可以由 FAHP 方法[21] 确定。

定义 9　设有两个云分别为 $T_1(Ex_1, En_1, He_1)$、$T_2(Ex_2, En_2, He_2)$，记 $T = T_1 \oplus T_2$ 为 T_1 与 T_2 的合成，则云的加权合成公式如下所示:

$$T(Ex, En, He) = (\lambda_1 \times T_1) \oplus (\lambda_2 \times T_2) \oplus \cdots \oplus (\lambda_m \times T_m) \tag{11}$$

其中:

$$Ex = \sum_{i=1}^{m} (\lambda_i \times Ex_i)$$

$$En = \sqrt{\sum_{i=1}^{m} (\lambda_i \times En_i^2)}$$

$$He = \sum_{i=1}^{m} (\lambda_i \times He_i)$$

2.8　信任云的决策

信任决策就是以两个信任云的比较结果为依据，选择满足预设条件的云的过程。由于信任云通过一维正态云刻画信任度值，故不能用曲线拟合等传统方法进行两个云的对比。因此，本文利用云的不同形态表征不同信任语意的特性，借助云相似度，实现云的比较。

定义 10　设 $TC_1(Ex_1, En_1, He_1)$ 和 $TC_2(Ex_2, En_2, He_2)$ 为两个信任评价云。将云 TC_1 经过正向云生成器生成云滴 (x_i, μ_i)，设 x_i 在云 TC_2 中的隶属度是 μ'_i，则称 $\frac{1}{n}\sum_{i=1}^{n} \mu'_i$ 为云 TC_1 和 TC_2 的相似度，记为 η。

云相似度计算算法如算法 3 所示。

算法 3　云相似度计算算法。

输入　云 $TC_1(Ex_1, En_1, He_1)$，云 $TC_2(Ex_2, En_2, He_2)$，N。

输出　相似度 K。

```
for i = 1:N                                   //N 个云滴
  Enn = normrnd(En1,He1);
// 云 TC1 中产生一个以 En1 为期望,He1 为标准差的正态随机数
  Enn = abs(Enn);
  x(i) = normrnd(Ex1,Enn);
// 云 TC1 中产生一个以 Ex1 为期望,Enn 为标准差的正态随机数
  Enm = normrnd(En2,He2);
// 云 TC2 中产生一个以 En2 为期望,He2 为标准差的正态随机数
  Enm = abs(Enm);
  if Enm ~ = 0                    // 将 xi 代入云 TC2 的期望方程中
    y(i) = exp(-(x(i) - Ex2)^2/(2 * Enm^2));
  else
    y(i) = 1;
  end
End
K = mean(y)                                   // 求得相似度
```

定义 11　正态云的 $3En$ 规则。在信任云的模型中，对于论域 U 中所有对其有贡献的云滴主要落在 $[Ex - 3En, Ex + 3En]$ 中，落在此区间以外的云滴的发生是小概率事件，可以忽略其贡献。

为得到可信节点，需通过信任决策方法将综合信任云与基准云实行全面的比较分析得出结论。信任云决策算法见算法 4。在云的三个特征参数中，反映知识中心值的期望值和反映分布裕度的熵直接影响了云形态的呈现。超熵反映熵的不确定性，只要超熵和熵的比值在合理的范围内，可以不考虑超熵参数的影响。信任决策分析就是对期望值和熵分析。在理想的情况下，信任评价云的期望值越大，分布裕度越小，该信任

云所表示的节点受信程度越高且信任值越高。下面针对决策分析时可能遇到的四类云图形进行讨论:

设 $TC(Exc,Enc,Hec)$ 为信任评价云(比较云),$TS(Exs,Ens,Hes)$ 为预设好阈值的基准云。

1) 当 $Exc > Exs$,$Enc < Ens$ 时,此比较云可信,属于较理想状态。

2) 当 $Exc > Exs$,$Enc > Ens$ 时,此时需考虑云的分布裕度,需引入正态云的 $3En$ 规则进行判断。分别计算两个云的 $Ex-3En$ 值,若比较云的 $Ex-3En$ 值较大,在云图上分布在基准云右侧,比较云可信;若比较云的 $Ex-3En$ 值较小,此种情况比较复杂,需以判断两个云的相似度方式确定可信度。

3) 当 $Exc < Exs$,$Enc < Ens$ 时,这种情况也需要引入正态云的 $3En$ 规则进行判断。若比较云的 $Ex-3En$ 值较大,需以判断两个云的相似度方式确定可信度;若比较云的 $Ex-3En$ 值较小,在云图上分布较为左侧,比较云不可信。

4) 当 $Exc < Exs$,$Enc > Ens$ 时,比较云的信任度值较低,不可信。

算法4 信任云决策算法。

输入 比较云 $TC(Exc,Enc,Hec)$,基准云 $TS(Exs,Ens,Hes)$。

输出 可信或不可信。

```
If (Enc/Hec < 5) return 0          // 不可信,比较云的离散度较大
If (Exc ≥ Exs)
  If (Enc ≤ Ens)   return 1  // 可信,表明期望值高,分布范围小
    else if (Exc - 3Enc ≥ Exs - 3Ens) return 1
                              // 可信,比较云整体信任值大于基准
    else  call 算法3                  // 使用云相似度计算算法
  Else
  If (Enc > Ens) return 0                              // 不可信
  else if (Exc - 3Enc ≥ Exs - 3Ens)
        call 算法3
  else return 0                    //不可信,比较云整体信任值较小
```

3 理论分析与性能评价

理论分析部分首先验证本文所提的节点信任评估机制是否满足预设目标;在此基础上,论述该机制的优点及合理性,从评估指标的全面性、算法复杂度方面进行分析并给出与同类型方案对比。性能评价部分是基于实验结果分析对本文方法的有效性、信任计算方法的合理性进行阐述,验证方案的可行性。

3.1 理论分析

本文方案针对云数据自毁方案中的 DHT 网络节点建立信任评估机制,在兼顾该应用场景下对 DHT 节点的性能表现和交互表现考察的基础上,满足系统信任评估的功能需求。在信任评估过程中,综合考虑了直接信任值、交易频度、推荐信任值、时间衰减因素、惩罚因子、存储时间、节点性能表现,从而对节点的信任度进行一个全面、可靠的评价。方案首先提出将节点性能表现评估作为评估指标,旨在从节点内部因素角度提供评估依据。节点性能包括节点生存参数、流量参数、搜索性能等;若节点的性能状况较差,可能直接导致密钥分量丢失。为了保证信任值的实效性和合理性,在进行信任值计算时不是简单地对历史信任值进行累加求平均。在现实情况中,信息的价值随时间推移而降低,因此计算时也需考虑信任值的时间衰减影响。惩罚因子用于控制节点服务的失败率,当一个节点出现服务失败的情况时,通过降低信任值的方式对其进行惩罚、区别,将失败次数与信任值相关联确保系统决策时能够选出更可信的节点。同时节点的存储时间需要长久保存,有利于对历史信任值进行迭代计算,进一步提高评估准确性。在同类型方案中,文献[12]中提出的 DTrust 方案仅考虑了前两种影响因素,缺乏对推荐信任值、时间衰减因素、惩罚因素、节点性能等方面的考虑。下面给出上述两种方案的对比汇总,如表2所示。

表2 DHT 节点信任评估方案对比

影响因素	DTrust 方案[12]	本文方案
节点直接信任值	是	是
推荐信任值	否	是
时间衰减	否	是
交易频度	是	是
惩罚因子	否	是
节点性能表现	否	是
长时间存储	否	是

在算法复杂度方面,本文方案直接信任值计算、推荐信任值计算、综合信任值计算、决策算法的算法复杂度均为 $O(n)$。在 DTrust 方案中各算法的复杂度为 $O(n^2)$。因此,本文方案在算法复杂度上表现更优,对比情况如表3所示。

表3 方案算法复杂度对比

类别	DTrust 方案[12]	本文方案
直接信任值	$O(n^2)$	$O(n)$
推荐信任值	—	$O(n)$
综合信任值	$O(n^2)$	$O(n)$

3.2 性能分析

首先通过仿真实验进行验证,实验目的是验证本文提出的面向云数据安全自毁的 DHT 网络节点信任评估机制的有效性,即验证节点信任值计算方法和信任决策算法的可行性、合理性。实验的环境配置为 Intel Core i5 CPU,4 GB 内存,Window 7 旗舰版操作系统(64 位),应用软件使用 Matlab R2012b。

为简化本次实验,本文在考虑时间因素对信任值的影响时,假设 DHT 网络节点每次服务间隔时间为 5 h。在多次实验的基础上,本文对方案中设计的各项常数、平衡系数取经验值,设置如下:交易频量常数 $C=0.4$,时间衰减平衡系数 $k=0.000\,35$。方案中各级指标对应的权重设置如图3所示。

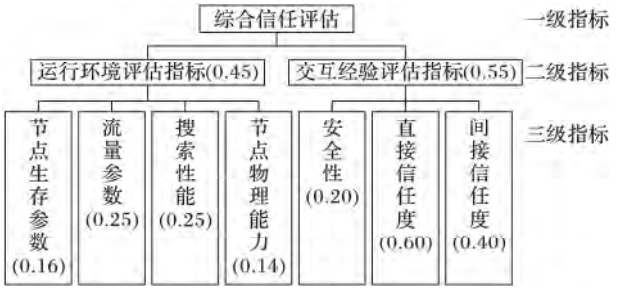


图3 信任机制中各级指标权重设置(数值为权重)

1) 数据预处理。

a) 给出运行环境 100 次评价向量,结合式(1)、式(11)并用逆向云生成器得到相应的三个特征参数,如表4所示。

b) 给出节点交互经验 100 次评价向量,并通过式(2) ~ (11) 计算其三个特征参数,如表 5 所示。

表 4　运行环境格式化数据

节点	$T(Ex,En,He)$
A	(0.892 8,0.182 9,0.018 3)
B	(0.737 3,0.174 1,0.016 9)
C	(0.604 2,0.183 8,0.018 1)
D	(0.311 4,0.176 7,0.017 5)
E	(0.095 8,0.152 4,0.014 4)

表 5　格式化数据

节点	$V(a,b,c,d,e)$	$T(Ex,En,He)$
A	V(80,10,10,0,0)	(0.925 0,0.160 1,0.016 1)
B	V(10,80,10,0,0)	(0.750 0,0.111 8,0.011 2)
C	V(20,20,60,0,0)	(0.650 0,0.200 0,0.020 1)
D	V(0,10,10,70,10)	(0.300 0,0.187 1,0.018 8)
E	V(10,10,20,20,40)	(0.325 0,0.336 3,0.033 8)

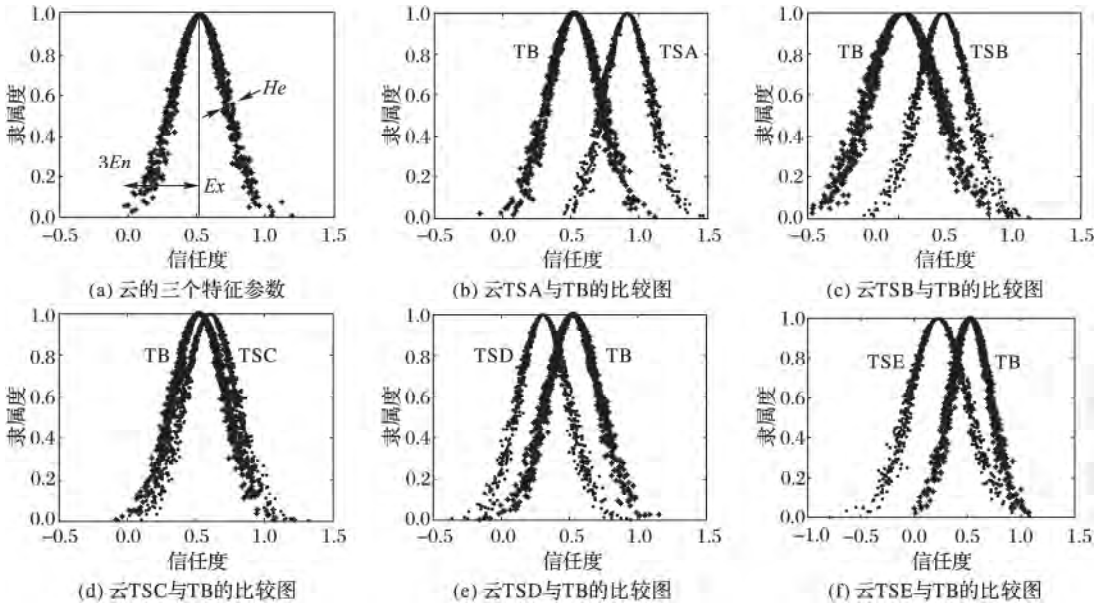
2) 假设信任基准云为 $TB(0.525\,0,0.192\,0,0.019\,3)$ 其对应信任向量为 $V(5,15,70,5,5)$。用式(11) 计算各个节点的综合评价云,并通过逆向云生成器生成云图形结果如图 4 所示(图中 TSA ~ TSE 为云名称)。

3) 实验结论与分析。

云的三个特征参数决定了云的形状,期望值 Ex 是云滴样本的均值,也被认为是节点的信任值,期望值越大在图形上分布越靠右侧; 熵 En 反映云滴的分布范围,熵越大云滴分布越集中,超熵 He 反映了云滴的离散程度,它们在图上的表现如图 4(a)。由信任云知识和信任决策算法可知,云的期望值越大,云滴分布的范围越小,则该云图形所代表的节点越可靠。

经过计算,通过 2.6 节提出的计算公式计算得到的信任度值符合正常预期: ① 信任值 $Trust$ 能匹配节点的行为表现,正确反映节点的信任级别,如节点 C 的信任值 $Trust_C$ = 0.610 9,对应级别为可信; 对节点 E,$Trust_E$ = 0.221 9,对应级别为不可信。② 当节点不诚信行为只有 1 次时,其不诚信行为造成信任值较小幅度的降低; 当节点不诚信行为达到 10 次时,对信任值影响较大。这是因为,为避免少数的不可信服务影响节点可信度,当某节点的不可信服务不超过 5 次时,惩罚值较小且随次数的增加而增大; 当某节点的不可信服务超过 5 次时,则认为该节点不可信任,惩罚值较大且随次数的增加而增大,当增加到一定次数时惩罚因子不变。图 4(b) 和(c) 中,TSA、TSB 的云图形分布在基准云图形的右侧,同时比较云的样本中心值大于基准云的样本中心值、比较云分布裕度较小(熵值决定云滴分布较集中),依据本文提出的信任决策算法,可判断节点 A、B 可信。图 4(f) 中,TSE 的云图形分布在左侧,特征参数 Ex、En 均小于基准云的对应值,故可判断节点 E 不可信。对于云 TSC、TSD 则需通过正态云"$3En$ 规则"和相似度才能作出进一步的判断。图 4(d) 中,云 TSC 分布在基准云右侧,熵值略大于基准云的对应值,计算出云 TSC 的 $Ex-3En$ 值大于基准云的 $Ex-3En$ 值,可判断节点 C 一般可信。图 4(e) 中,云 TSD 分布在基准云左侧且较分散,又由于 TSD 的 $Ex-3En$ 值小于基准云的 $Ex-3En$ 值,情况更为复杂,还需要进行相似度判断。本文给定相似度阈值为 0.5,借助相似度算法计算得云 TSD 的相似度为 0.49,小于给定阈值,因此,节点 D 不可信。

综上所述,信任决策能够判断节点的可信度等级,实验结果充分符合预期。面向云数据安全自毁的 DHT 网络节点信任评估机制的有效性得证。



(a) 云的三个特征参数　(b) 云TSA与TB的比较图　(c) 云TSB与TB的比较图

(d) 云TSC与TB的比较图　(e) 云TSD与TB的比较图　(f) 云TSE与TB的比较图

图 4　比较云与基准云

4　结语

随着人们迈入互联网 3.0 时代,云存储服务的出现改变了人们数据存储的方式,同时也导致用户对数据的不可控,进而引发一系列安全问题。在这种情况下,相关领域的学者借助 DHT 网络提出一种新的数据安全自毁方案以应对这种矛盾局面,但 DHT 网络本身存在不确定性,会给系统带来某些潜在威胁。为此,本文提出面向云数据安全自毁的 DHT 网络节点信任评估机制,首先用信任云模型对节点的信任实现定性定量的刻画; 同时,提出了新的计算节点直接信任值和间接信任值的方法,综合考虑节点的内外因素,细化到不同层级对节点进行信任评估,然后将各层评价子云加权合成得到反馈

行为信息的综合信任云;最后,通过信任决策算法选择可信节点。本文针对节点信任值计算方法和信任评估过程进行实验,验证了面向云数据安全自毁的DHT网络节点信任评估机制的有效性和可行性。该机制的提出,提高了系统的抗攻击性并降低了系统运算负载,但还存在一些不足,进一步研究信任影响因素以及移动端云数据信任评估方案[23]是下一步工作的重点。

参考文献:

[1] XIONG J B, LIU X M, YAO Z Q, et al. A secure data self-destructing scheme in cloud computing [J]. IEEE Transactions on Cloud Computing, 2014, 2(4): 448 - 458.

[2] XIONG J B, LI F H, MA J F, et al. A full lifecycle privacy protection scheme for sensitive data in cloud computing [J]. Peer-to-Peer Networking and Applications, 2015, 8(6): 1025 - 1037.

[3] PERLMAN R. File system design with assured delete [C] // Proceedings of the 3rd IEEE International Security in Storage Workshop. Piscataway, NJ: IEEE, 2005: 83 - 88.

[4] YANG T, LEE P P, LUI J C, et al. FADE: secure overlay cloud storage with file assured deletion [C] // Proceedings of the 2010 International Conference on Security and Privacy in Communication Systems. Berlin: Springer, 2010: 380 - 397.

[5] NAIR S K, DASHTI M T, CRISPO B, et al. A hybrid PKI-IBC base ephemerizer system [C] // Proceedings of the 2007 International Information Security Conference. Berlin: Springer, 2007: 241 - 252.

[6] GEAMBASU R, KOHNO T, LEVY A, et al. Vanish: increasing data privacy with self-destruction data [C] // Proceedings of the 18th USENIX Security Symposium. Berkeley: USENIX, 2009: 299 - 355.

[7] XIONG J B, YAO Z Q, MA J F, et al. A secure document self-destruction scheme with identity based encryption [C] // Proceedings of the 2013 5th International Conference on Intelligent Networking and Collaborative Systems. Piscataway, NJ: IEEE, 2013: 239 - 243.

[8] XIONG J B, YAO Z Q, MA J F, et al. A secure document self-destruction scheme: an ABE approach [C] // Proceedings of the 2013 IEEE 10th International Conference on High Performance Computing and Communications & 2013 IEEE International Conference on Embedded and Ubiquitous Computing. Piscataway, NJ: IEEE, 2013: 59 - 64.

[9] ZENG L, SHI Z, XU S, et al. SafeVanish: an improved data self-destruction for protecting data privacy [C] // Proceedings of the IEEE Second International Conference on Cloud Computing Technology and Science. Piscataway, NJ: IEEE, 2010: 521 - 528.

[10] WOLCHOCK S, HOLFMANN O S, HENINGER N, et al. Defeating vanish with low-cost sybil attacks against large DHTs [C] // Proceedings of the 17th Annual Network & Distributed System Security Conference. San Diego: [s. n.], 2010: 1 - 15.

[11] 熊金波,姚志强,马建峰,等. 面向网络内容隐私的基于身份加密的安全自毁方案[J]. 计算机学报, 2014, 37(1): 140 - 150. (XIONG J B, YAO Z Q, MA J F, et al. A secure self-destruction scheme with IBE for the Internet content privacy [J]. Chinese Journal of Computers, 2014, 37(1): 140 - 150.)

[12] 吴守才. 云环境中基于信任评估的数据自销毁机制研究[D]. 沈阳:辽宁大学, 2015: 1 - 60. (WU S C. Research on trust-based self-destructing scheme in cloud computing [D]. Shenyang: Liaoning University, 2015: 1 - 60.)

[13] 杨莎. 基于云理论的信任评估模型及应用研究[D]. 北京:华北电力大学, 2011: 1 - 53. (YANG S. Research on trust evaluation model and application based on cloud theory [D]. Beijing: North China Electric Power University, 2011: 1 - 53.)

[14] LI T L, ZHOU X B. ZHT: a light-weight reliable persistent dynamic scalable zero-hop distributed Hash table [C] // Proceedings of IEEE 27th International Parallel Distributed Processing Symposium. Piscataway, NJ: IEEE, 2013: 775 - 787.

[15] FALKER J, POATEK M, JOHN P, et al. Profiling a million user DHT [C] // Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement. New York: ACM, 2007: 129 - 134.

[16] 翟学伟. 信任的本质及其文化[J]. 社会, 2014, 34(1): 1 - 26. (ZHAI X W. The Essence of trust and its culture [J]. Chinese Journal of Sociology, 2014, 34(1): 1 - 26.)

[17] 魏达,贾翔鹏,王健,等. 基于可信证书的可信网络接入模型及实现[J]. 吉林大学学报(工学版), 2010, 40(2): 496 - 500. (WEI D, JIA X P, WANG J, et al. New access model and implementation of trusted network based on trusted certificate [J]. Journal of Jilin University (Engineering and Technology Edition), 2010, 40(2): 496 - 500.)

[18] 张仕斌,许春香. 基于云模型的信任评估方法研究[J]. 计算机学报, 2013, 36(2): 422 - 430. (ZHANG S B, XU C X. Study on the trust evaluation approach base on cloud model [J]. Chinese Journal of Computers, 2013, 36(2): 422 - 430.)

[19] 路峰,吴慧中. 网格环境下基于云模型的信任评估和决策方法研究[J]. 系统仿真学报, 2009, 21(2): 421 - 426. (LU F, WU H Z. Research of trust valuation and decision-making based on cloud model in grid environment [J]. Journal of System Simulation, 2009, 21(2): 421 - 426.)

[20] 李德毅,孟海军,史雪梅. 隶属云和隶属云发生器[J]. 计算机研究与发展, 1995, 32(6): 15 - 20. (LI D Y, MENG H J, SHI X M. Membership clouds and membership cloud generators [J]. Journal of Computer Research and Development, 1995, 32(6): 15 - 20.)

[21] 王亮,郭亚军. P2P系统中基于声誉的信任评估机制[J]. 计算机工程与应用, 2009, 45(15): 136 - 138. (WANG L, GUO Y J. Reputation-based on trust evaluation mechanism for P2P system [J]. Computer Engineering and Applications, 2009, 45(15): 136 - 138.)

[22] 章夏彦,张少敏. 信任计算在大用户直购电交易中的研究与应用[D]. 北京:华北电力大学, 2010: 1 - 55. (ZHANG X Y, ZHANG S M. Research and application of trust computing in direct power power purchase for large consumers [D]. Beijing: North China Electric Power University, 2010: 1 - 55.)

[23] 沈薇薇,姚志强,熊金波,等. 面向移动终端的隐私数据安全存储及自毁方案[J]. 计算机应用, 2015, 35(1): 77 - 82. (SHEN W W, YAO Z Q, XIONG J B, et al. Secure storage and self-destruction scheme for privacy data in mobile devices [J]. Journal of Computer Applications, 2015, 35(1): 77 - 82.)

Background

This work is partially supported by the National Natural Science Foundation of China (61402109, 61370078), the Natural Science Foundation of Fujian Province (2015J05120), the Distinguished Young Scientific Research Talents Plan in Universities of Fujian Province (2015).

WANG Dong, born in 1993, M. S. candidate. His research interests include trust evaluation, data security.

XIONG Jinbo, born in 1981, Ph. D., associate professor. His research interests include cloud data security, privacy protection.

ZHANG Xiaoying, born in 1995. Her research interests include trust evaluation, data security.