

云计算环境下基于用户行为信任评价研究

谭跃生, 王 超

(内蒙古科技大学 信息工程学院, 内蒙古 包头 014010)

摘 要: 针对云计算开放环境下的安全问题, 为有效分析用户不可信行为和异常行为, 实现用户行为信任的科学量化评估, 综合考虑信任评估的动态性, 提出一种基于相对熵的组合赋权法用户行为信任评估, 结合主客观赋权法的权重信息, 通过相对熵计算各主客观赋权法权重的贴近度, 从而确定各赋权法权重在组合权重中的加权系数, 最后对组合权重运用模糊综合评价得到用户的信用度。验证结果实现了模型的具体化, 表明了组合赋权法的科学性和可操作性, 并对用户的信用度做出了客观评价。

关键词: 云计算; 用户行为; 相对熵; 组合权重

中图分类号: TP391

文献标识码: A

文章编号: 1000-7180(2015)11-0147-05

Trust Evaluation Based on User Behavior in Cloud Computing

TAN Yue-sheng, WANG Chao

(School of Information Engineering, Inner Mongolia University of Science and Technology,
Baotou 0140010, China)

Abstract: Aim at the security issues in cloud computing, a method based on Combination Weighting Approach of Relative Entropy is proposed to evaluate user's behavior trust to analyze the untrust and abnormal behavior effectively, so as to achieve the scientific quantitative evaluation for user's behavior trust. Combined with the Weighting factor of weighting method which is both subjective and objective, relative entropy is applied to compute the approach degree of weights. Which are set by subjective and objective weighting method, thus the weighting efficient of each weighting method's weight in combination weight can be determined. Finally, fuzzy comprehensive assessment is made for combination weight to get user's trust value. Implementation and verification of the model shows that combination weighting method and the model are scientific and operability, and making objective evaluation for the user's credit.

Key words: cloud computing user's behavior; relative entropy; combination weighting

1 引言

云计算将海量的数据资源链接在一起, 构成了一个超大规模的共享虚拟资源池, 极大方便了云端用户使用资源, 并且降低了云端用户使用存储资源消耗的成本, 但这也给一些非法用户破坏云端资源提供了可趁之机。正是这些显著的优势和安全隐患, 使得云计算环境下多租户访问过程中的安全问题逐步成为人们关注并研究的课题。今天, 对用户身份真

实性验证的同时, 也要对用户的行为进行信任评估。身份验证技术虽然已经比较成熟, 仅用身份验证并不能完全阻止恶意攻击对云资源的破坏, 因此对用户行为进行分析评估是有效提升云安全的关键, 也是今天云计算研究中的一个重点。

近年来, 各国学者对信任机制展开了研究。文献[1]中, 在云计算环境下, 利用基于三角模糊数的模糊网络分析法(FANP)对网络用户行为属性的权重进行量化计算, 并对用户行为进行信任评, 但该文献

收稿日期: 2015-01-20; 修回日期: 2015-03-09

基金项目: 国家自然科学基金(61462069); 内蒙古自然科学基金(2012ms0912, 2014ms0622)

利用单一的模糊分析法计算权重,其权重的主观性还很大程度的存在.文献[2]中,在云计算环境下提出了一种新的基于三角模糊层次分析法(FAHP)的用户行为信任评估方法来确定用户各行为指标的权重,实现了用户行为信任的科学量化信任评估,但此方法和文献[1]中的一样,都不能在一定程度上降低其确定权重的主观性.文献[3]中提出一种多维决策属性的信任关系量化模型,模型中加入了一系列的决策属性,如风险函数、激励函数等,利用熵权法计算各个决策属性的权重,弱化了计算权重的主观性,使得此模型具有更加稳健的动态适应性,模型的安全性有明显加强.文献[4]中针对云计算环境下频繁出现的恶意攻击行为,提出了一种信任防御模型,该模型在收集用户行为证据后,选用模糊层次分析法(FAHP)对云端用户行为进行评估,从而确定各行为的权重,它的不足之处在于权重方法的选择不可避免具有人为主观性.文献[5]中提出了一种用户行为评估层次结构模型,该模型所选用的评估方法能够比较准确的确定用户行为证据的权重,但该文献在计算用户权重时使用的是主观赋权方法 AHP,权重的主观性没有减弱.文献[6]中提出了一种适用于网格环境基于 Fussy Logic 的信任评估模型,该模型具有较好的恶意实体检测、抵御能力;不足在于其收敛性较低、未评估间接信任度,全局可信性较低.文献[7]中在普适计算环境下提出了一种基于不同实体间向量运算机制的信任模型,其信任评估取决于对等实体共同的交互实体的推荐,将历史、信任和时间等影响因子引入来反映用户信任评估的动态性,但其没有解决交互实体推荐时的欺骗行为.

本文根据上述研究,将基于相对熵的组合赋权法应用到用户行为信任评估中,通过相对熵计算得出各主客观赋权法权重的贴近度,根据各主客观赋权法权重的贴近度确定各赋权法权重在组合权重的加权系数.最后对组合权重进行模糊综合评价,从而对用户信用度进行评价.

2 用户信任评估

2.1 相关定义

定义 1 信任是指信任的主体对信任客体行为的能力、诚实度、安全性和可靠性的主观信念,这种信念依赖于特定的上下文^[8].

定义 2 用户行为证据是指借助计算机软硬件检测获取的基础数值,这些基础数据用来定量评价用户的总体行为.

定义 3^[9] 如果 $x_i, y_i > 0, i = 1, 2, 3, \dots, n$ 且

$$1 = \sum_{i=1}^n x_i > \sum_{i=1}^n y_i, \text{ 称 } H(x, y) = \sum_{i=1}^n x_i \log \frac{x_i}{y_i} \text{ 为 } x$$

相对于 y 的相对熵.具有的主要性质如下:

$$(1) \sum_{i=1}^n x_i \log \frac{x_i}{y_i} \geq 0;$$

$$(2) \sum_{i=1}^n x_i \log \frac{x_i}{y_i} = 0 \text{ 的充分必要条件是 } x_i = y_i > 0, i = 1, 2, \dots, n.$$

定义 4 信任度可以称作一个实体对另一个实体的信任级别,也就是信任的程度^[10].

2.2 信任评估模型

按照实际应用中的需求和功能特性将整体的用户行为信任进行分层分解,将综合、不明确的用户行为信任分解为若干行为信任属性,这样可以有效解决云计算环境下用户行为信任的笼统性和不确定性问题^[11].将用户行为信任层次模型从高到底分解成目标层(O)、准则层(C)和指标层(I).如图 1 所示.

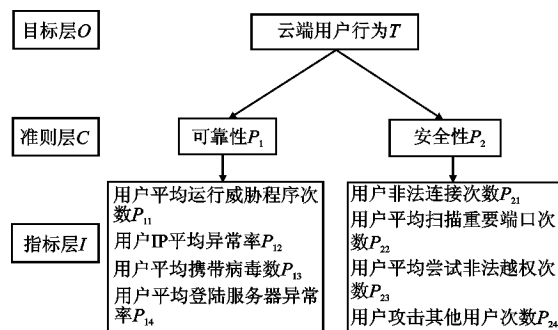


图 1 用户行为信任层次模型

目标即要解决问题的目的,也是模型要实现的总目标.准则层是采用某种措施、方案等实现预期总目标所设置的中间层次.指标层表示其上层即准则层中选用的解决问题的各种措施和方案等,通常可有若干个方案供选择.将模型用一个三元组可以表示成: $M = \{O, C, I\}$, 其中 $C = \{C_1 C_2 \dots C_n\}$, $C_1 C_2 \dots C_n$ 是准则层的子准则且 $C_i \cap C_j = \emptyset (i \neq j, 1 \leq i, j \leq n)$. $I = \{I_1 I_2 \dots I_p \dots I_n\}$, $I_1 I_2 \dots I_p \dots I_n$ 是指标层 I 中的策略.

为了验证组合权重的可行性,本文选取模糊综合评价模型对用户信任进行综合评价.给定两个有限论域: $U = \{u_1, u_2, \dots, u_n\}$ 和 $V = \{v_1, v_2, \dots, v_n\}$.其中, U 代表用户行为属性组成的集合, V 代表信任度组成的集合.

用 A 表示第一层属性的权重, R 表示用户信用评价的单属性模糊评价矩阵, R_m 表示第 m 个单属性

评价结果, B 表示用户信任评价的第一层模糊综合评价集. 本文将每个用户行为信任属性的信任度分为 5 等, 分别为信任、较信任、一般信任、较不信任和不信任.

模型的关键在于构造单属性模糊评价矩阵. 根据最大隶属度原则, 设单属性模糊隶属度为 r_{ij} , 构造出的单属性模糊评价矩阵 R 为

$$R = \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_m \end{bmatrix} = \begin{bmatrix} r_{11} & \cdots & r_{15} \\ \vdots & \ddots & \vdots \\ r_{m1} & \cdots & r_{m5} \end{bmatrix} \quad (1)$$

若各评价属性的权重集为 $A = [a_1, a_2, \cdots, a_m]$, 应用乘法运算, 可得到集合 V 上的一个模糊子集, 即综合评价结果:

$$B = AR = (b_1, b_2, \cdots, b_n) \quad (2)$$

2.3 行为证据规范化

用户行为证据表现形式有多种, 常见的如下所示.

- (1) 在一个大体范围内的具体数值.
- (2) 百分比形式.

证据数据的表现形式多种多样, 为了便于计算和评估, 将证据数据进行无量纲化处理, 因此, 要对数据进行规范化, 将这些数据规范到 $[0, 1]$ 区间内.

百分比表示的证据本身就在 $[0, 1]$ 范围内, 只需通过式(3) 将其转化为正向递增数值:

$$e_{ij} = \begin{cases} a_{ij}, a_{ij} \text{ 正向递增} \\ 1 - a_{ij}, a_{ij} \text{ 正向递减} \end{cases} \quad (3)$$

式中, a_{ij} 为初始数据, e_{ij} 为规范化的数据.

具体的数值如果不在 $[0, 1]$ 区间内, 可以通过式(4) 将其规范化:

$$e_{ij} = \begin{cases} \frac{a_{ij} - (a_{ij})_{\min}}{(a_{ij})_{\max} - (a_{ij})_{\min}}, a_{ij} \text{ 正向递增} \\ \frac{(a_{ij})_{\max} - a_{ij}}{(a_{ij})_{\max} - (a_{ij})_{\min}}, a_{ij} \text{ 正向递减} \end{cases} \quad (4)$$

式中, $(a_{ij})_{\min}$ 和 $(a_{ij})_{\max}$ 分别为初始证据的最小值和最大值.

2.4 组合赋权法

赋权方法的种类繁多, 我们不能绝对地认为某一种赋权方法计算得出的权重是准确的, 也不能说它比其他的赋权方法合理. 选择相对科学合理的赋权方法在一定程度上会影响评价结果. 为此我们提出了基于相对熵的组合赋权法, 在使用组合赋权法时, 若能确定选用的每一种主客或客观赋权法的“可信度”, 我们就可以依据此可信度对不同赋权方法的

权重进行加权平均. 本文根据相对熵对所采用的主、客观赋权法权重进行集结得到集结权重, 利用主、客观赋权法权重与集结权重之间的贴近度度量主、客观赋权法权重的可信度.

α_k 的确定如下所示.

设 μ_k 为选用的赋权方法里的一个权向量:

$$\mu_k = (\mu_{k1}, \mu_{k2}, \cdots, \mu_{kn}) \quad (k = 1, 2, \cdots, q) \quad (5)$$

式中, ω_j 为组合权重, α_k 表示第 k 个权向量的可信度, 且

$$\sum_{k=1}^q \alpha_k = 1, \text{ 那么:}$$

$$\omega_j = \sum_{k=1}^q \alpha_k \mu_{kj}, j \in n \quad (6)$$

定义任意两种权重方法的权向量: $u_i, u_j (i, j = 1, 2, \cdots, p)$, 它们之间的相对熵可以表示为

$$H(\mu_i, u_j) = \sum_{l=1}^m u_{jl} \log \frac{u_{il}}{u_{jl}} \quad (7)$$

由相对熵的定义可知 $H(\mu_i, u_j) = 0$ 当且仅当 $\forall l \in n \exists u_{il} = u_{jl}$. 因此, 相对熵 $H(u_i, u_j)$ 可用于度量两种不同赋权方法所得权向量 μ_i, μ_j 的贴近程度, 即贴近度. 对各种赋权法权重结果进行集结得到的组合权重可以通过数学规划模型来求解, 如下:

$$\begin{cases} \min H(v) = \sum_{j=1}^q \sum_{i=1}^m v_i \log \frac{v_i}{u_{ji}} \\ \text{s. t. } \sum_{i=1}^m u_i = 1, u_i > 0 \end{cases}$$

该模型有全局最优解 $\omega = (\omega_1, \omega_2, \cdots, \omega_m)$, 其中,

$$v_i = \frac{\prod_{j=1}^q (u_{ji})^{\frac{1}{q}}}{\sum_{i=1}^m \prod_{j=1}^q (u_{ji})^{\frac{1}{q}}} \quad (8)$$

根据模型的全局最优解可以计算得出基于相对熵的组集结权重^[12], 下面给出计算第 k 个赋权法权重的可信度 α_k 的步骤.

步骤 1 从主观和客观赋权法中选取 q 种不同赋权方法, 计算 q 种赋权方法的权重向量:

$$\mu_k = (\mu_{k1}, \mu_{k2}, \cdots, \mu_{kn}) \quad (k = 1, 2, \cdots, q) \quad (9)$$

步骤 2 根据式(5) 和式(7) 建立最优化模型, 利用式(8) 得出组合权重 $\omega = (\omega_1, \omega_2, \cdots, \omega_m)$.

步骤 3 计算每一种权重和组合权重的贴近度:

$$h(\mu_i, \omega) (i = 1, 2, \cdots, q) \quad (10)$$

步骤4 根据步骤3中的贴进度,计算 q 种赋权方法权重的可信度.第 k 个赋权法权重与组合权重的贴进度值越大,则其在组合赋权中的作用也就越大^[13],它的可信度可以表示为

$$\alpha_i = \frac{h(\mu_i, \omega)}{\sum_{i=1}^q h(\mu_i, \omega)} \quad (11)$$

有上述步骤可以计算出各赋权法权重的可信度

α_i ,代入公式 $\omega_j = \sum_{k=1}^q \alpha_k \mu_{kj}$ 中可以算出组合权重.

3 模型实现与验证

从科学性、层次性和可操作性的角度出发,本文从可靠性和安全性两个方面来构建用户行为信任层次模型^[14],通过仿真实验来实现和验证模型,仿真实验中随机选取5个用户.这5个用户的8个用户行为指标详见图1.

3.1 各个赋权方法权重计算结果

首先将用户行为证据标准化,标准化后的数据如表1所示.

表1 标准化数据

	用户1	用户2	用户3	用户4	用户5
P_{11}	0.428 6	0.285 8	1.000 0	0.000 0	0.714 3
P_{12}	0.666 7	0.333 3	1.000 0	0.000 0	0.666 7
P_{13}	0.769 2	0.000 0	0.846 2	0.230 8	1.000 0
P_{14}	0.333 3	0.666 7	1.000 0	0.333 30	0.000 0
P_{21}	0.666 7	0.555 6	1.000 0	0.000 0	0.777 8
P_{22}	0.428 6	0.428 6	1.000 0	0.000 0	0.714 6
P_{23}	0.428 6	0.000 0	0.428 6	0.142 9	1.000 0
P_{24}	0.666 7	0.333 3	1.000 0	0.000 0	0.833 3

分别采用AHP法、ANP法、熵权法、变异系数法对上述用户行为指标进行赋权,通过仿真实验所得到的权重如表2所示.

表2 评价指标权重

	AHP法	ANP法	熵权法	变异系数法	集结权重
P_{11}	0.065 8	0.183 6	0.163 0	0.130 8	0.136 7
P_{12}	0.035 3	0.066 6	0.110 4	0.117 6	0.080 2
P_{13}	0.113 7	0.143 1	0.123 9	0.118 9	0.134 3
P_{14}	0.035 3	0.106 7	0.127 1	0.134 1	0.096 6
P_{21}	0.274 2	0.135 4	0.094 4	0.102 5	0.166 5
P_{22}	0.124 6	0.093 1	0.110 5	0.119 2	0.120 0
P_{23}	0.186 2	0.073 3	0.158 4	0.157 8	0.148 6
P_{24}	0.065 0	0.194 2	0.112 4	0.119 1	0.123 0

将上表中的各赋权法的权重代入式(8),求得各

指标的集结权重:

$$\begin{aligned} \omega_1 &= 0.136\ 7, \omega_2 = 0.080\ 2, \omega_3 = 0.134\ 3, \omega_4 \\ &= 0.096\ 6, \omega_5 = 0.166\ 5, \omega_6 = 0.120\ 0, \omega_7 = \\ &0.148\ 6, \omega_8 = 0.120\ 3. \end{aligned}$$

各种赋权法的权重与集结权重 ω 的贴进度 $h(\mu_i, \omega)$, ($i = 1, 2, 3, 4$):

$$\begin{aligned} h(\mu_1, \omega) &= 0.010\ 5, h(\mu_2, \omega) = 0.046\ 7, h(\mu_3, \\ \omega) &= 0.026\ 2, h(\mu_4, \omega) = 0.023\ 9. \end{aligned}$$

利用式(11)计算可信度 α_k ($k = 1, 2, 3, 4$):

$$\alpha_1 = 0.097\ 9, \alpha_2 = 0.477\ 0,$$

$$\alpha_3 = 0.244\ 2, \alpha_4 = 0.222\ 7.$$

再由式(6)计算得到组合权重 ω ,即:

$$\begin{aligned} \omega &= \\ & (0.163\ 0, 0.088\ 4, 0.136\ 1, 0.115\ 3, 0.137\ 3, 0.110\ 1, 0.127\ 0, 0.153\ 0). \end{aligned}$$

3.2 模糊综合信任评价

首先要建立模糊隶属度函数,本文采用模糊隶属度三角函数.构造完模糊隶属度三角函数后,将用户行为信任层次中每一个指标转换成一个模糊隶属度,构成用户信用模糊评价的单因素评价集.以用户平均携带病毒数 P_3 为例构建模糊隶属度三角函数,评语等级如表3所示.

表3 评语等级

不信任	较不信任	一般信任	较信任	信任
≥ 20	$[15, 20)$	$[10, 15)$	$(5, 10)$	$[0, 5]$

以用户平均扫描重要端口次数 P_{21} 为例构建模糊隶属度三角函数:

$$f_1(P_{21}) = \begin{cases} 0, (0, 18] \\ \frac{P_{13} - 18}{20 - 18}, (18, 20) \\ 1[20, +\infty) \end{cases} \quad (12)$$

$$f_2(P_{21}) = \begin{cases} \frac{P_{13} - 13}{18 - 13}, [13, 18) \\ \frac{20 - P_{13}}{20 - 18}, [18, 20) \\ 0, \text{其他} \end{cases} \quad (13)$$

$$f_3(P_{21}) = \begin{cases} \frac{P_{13} - 8}{13 - 8}, [8, 13) \\ \frac{18 - P_{13}}{18 - 13}, [13, 18) \\ 0, \text{其他} \end{cases} \quad (14)$$

$$f_4(P_{21}) = \begin{cases} \frac{P_{13} - 5}{13 - 8}, [5, 8) \\ \frac{13 - P_{13}}{13 - 8}, [8, 13) \\ 0, \text{其他} \end{cases} \quad (15)$$

$$f_5(P_{21}) = \begin{cases} 1 & [0, 5) \\ \frac{8-P_{13}}{8-5}, & [5, 8) \\ 0, & \text{其他} \end{cases} \quad (16)$$

根据构造的模糊隶属度三角函数计算得到用户 1 的单因素模糊隶属度如下:

$P_{11}: (0, 0, 0.33, 0.67, 0);$

$P_{12}: (0, 0, 1, 0, 0);$

$P_{13}: (0, 0.4, 0.6, 0, 0);$

$P_{14}: (0, 0, 0, 1, 0);$

$P_{21}: (0, 0, 0.67, 0.33, 0);$

$P_{22}: (0, 0, 0, 0, 1);$

$P_{23}: (0, 0, 0, 0, 1);$

$P_{24}: (0, 0, 0, 1, 0).$

再根据计算出的组合权重和公式 $B = AR = (b_1, b_2, \dots, b_n)$ 可以得到组合赋权法的用户信任模糊综合评价向量:

$B_1 = (0, 0.054\ 4, 0.415\ 9, 0.269\ 8, 0.227\ 1).$

根据最大隶属度原则,用户 1 的信任度为一般信任.以同样的方法可以得到其他 4 名用户的模糊综合评价向量和信任,如表 4 所示.

表 4 组合赋权法模糊综合评价结果

用户	对各级别的隶属度				
	不信任	较不信任	一般信任	较信任	信任
用户 1	0	0.054 4	0.415 9	0.269 8	0.227 1
用户 2	0	0.096 6	0.053 0	0.324 5	0.321 9
用户 3	0	0.337 4	0.450 8	0.040 6	0.148 6
用户 4	0	0	0.096 6	0.134 3	0.769 1
用户 5	0	0.134 3	0.507 1	0.359 0	0

从表 4 可以看出用户 1 的信任级别为一般信任,用户 2 的信任度为较信任,用户 3 的信任度为一般信任,用户 4 的信任度为信任,用户 5 的信任度为一般信任.

4 结束语

云计算的发展突飞猛进,面对其开放的运行环境,有效地对云端用户进行信任评估是重中之重,也是保证安全的基础.本文在建立信任评价模型的同时,综合主客观赋权法的权重信息提出了基于相对熵的组合赋权法,利用一个最优化数学模型求出组合赋权加权系数,避免了主观性,实验部分表明了组合赋权法的科学性和可操作性,模糊综合评价对用户行为信任做出了客观评价.但本文使用的组合赋权法是一种新型的系统赋权方法,其在用户行为信任评价中的应用还需要进一步改进.

参考文献:

- [1] 吕艳霞,田立勤,孙珊珊.云计算环境下基于 FANP 的用户行为的可信评估与控制分析[J]. 计算机科学, 2013,40(1):132-135.
- [2] 郭树凯,田立勤,沈学利. FAHP 在用户行为信任评价中的研究[J]. 计算机工程与应用, 2011, 47(12): 59-61.
- [3] 李小勇,桂小林. 可信网络中基于多维决策属性的信任量化模型[J]. 计算机学报, 2009,32(3):405-416.
- [4] 周茜,于炯. 云计算下基于信任的防御系统模型[J]. 计算机应用, 2011,31(6):1531-1535.
- [5] 冀铁果,田立勤,胡志兴,等. 可信网络中一种基于 AHP 的用户行为评估方法[J]. 计算机工程与应用, 2007,43(19):123-126.
- [6] Song SS, Hwang K. Fuzzy trust integration for security enforcement in grid computing[C] // Proc of the Int' 1 Symp on Network and Parallel Computing (NPC2004). Berlin:Springer-Verlag, 2005:9-21
- [7] Jameel H, Hung LX. A trust model for ubiquitous systems based on vectors of trust values[C] // Proc of the 7th IEEE Int'1 Symp on Multimedia. Washington: IEEEComputer Society Press, 2005: 674-679.
- [8] 周茜,于炯. 可信网络中基于信誉和风险评估的动态信任模型[J]. 计算机应用研究, 2010, 27(11): 4211-4214
- [9] 刘仁,卞树檀,谭营. 相对熵在多指标系统评估中的应用[J]. 四川兵工学报, 2012,33(5):116-118.
- [10] 李小勇,桂小林. 大规模分布式环境下动态信任模型研究[J]. 软件学报, 2007,18(6):1510-1521.
- [11] Tian L Q, Lin C, i T G. Kind of quantitative evaluation of user behavior trust using AHP[J]. Journal of Computational Information Systems, 2007, 3(4): 1329-1334.
- [12] 刘靖旭,谭跃进,蔡怀平. 多属性决策中的线性组合赋权方法研究[J]. 国防科技大学学报, 2005, 27(4): 121-124.
- [13] 周宇峰,魏法杰. 基于相对熵的多属性决策组合赋权方法[J]. 运筹与管理, 2006,15(5):48-53.
- [14] 查财旺. 云计算环境下行为信任模型研究[D]. 大连: 大连海事大学, 2011.

作者简介:

谭跃生 男,(1959-),教授.研究方向为云计算与信息安全. E-mail:tanys@imust.cn.

王超 男,(1987-),硕士研究生.研究方向为云计算与信息安全.