

# 基于 Shapley 熵的主观信任模型\*

刘宇轩, 郭玉翠

(北京邮电大学 理学院, 北京 100876)

**摘要:** 针对开放网络信任的主观性、不确定性等特点, 建立了一种基于 Shapley 熵的主观模糊信任综合评价模型。该模型运用模糊综合评价法得出节点间的直接信任值, 并提出评价取向相似度、反馈评价不确定度、推荐节点可信度三个概念, 利用 Shapley 熵计算节点推荐信任值。仿真实验表明, 该模型能准确地计算信任值, 并在保证节点交易成功率的同时体现了评价的主观性。

**关键词:** 信任管理; Shapley 熵; 评价取向相似度; 反馈评价不确定度; 推荐节点可信度; 模糊综合评判

**中图分类号:** TN918.4      **文献标志码:** A      **文章编号:** 1001-3695(2012)12-4656-03

doi:10.3969/j.issn.1001-3695.2012.12.065

## Trust management model based on Shapley entropy

LIU Yu-xuan, GUO Yu-cui

(Dept. of Mathematics, Beijing University of Posts & Telecommunications, Beijing 100876, China)

**Abstract:** Aiming at the features of subjectivity and uncertainty of trust in open network, this paper proposed a new subject comprehensive trust management model based on Shapley entropy. This model computed the direct trust value using fuzzy comprehensive evaluation method, and introduced the conception of Shapley entropy to calculate the recommendation trust value based on similarity of evaluating tendency, uncertainty of evaluation feedback and reliability of recommendation nodes, in order to manifest the subjectivity of evaluation while guaranteeing the security of transactions. Simulation results show that the model can calculate trust value accurately and ensure the success rate of interaction effectively.

**Key words:** trust management; Shapley entropy; similarity of evaluating tendency; uncertainty of evaluation feedback; reliability of recommendation nodes; fuzzy comprehensive evaluation

随着网络技术的迅速发展, P2P 网络因其开放性、匿名性和无中心性等特性得到广泛应用, 然而这些特性不仅使得网络的可靠性和安全性难以保证, 而且决定了完全的客观信任机制不能全面地反映实体间的信任关系, 所以主观信任关系正日益成为 P2P 网络安全研究中的焦点<sup>[1-9]</sup>。

1996 年, Blaze 等人<sup>[10]</sup>首先提出信任管理的概念, Policy-Maker 和 KeyNote 开发了两个早期的管理系统, 这些系统要求资源请求方单方面暴露所持证书, 会导致资源请求方的隐私泄露。之后比较有代表意义的有 Beth 模型<sup>[11]</sup>和 Jøsang 模型<sup>[12]</sup>等。Beth 模型基于概率模型对主观信任进行建模, 其不足之处在于没有考虑信任的模糊性, 而是将这种模糊性简单地等同于随机性。Jøsang 提出一种基于主观逻辑的信任模型, 引入证据空间和观念空间来描述信任关系, 但该模型实际上将信任的模糊性与随机性等同, 而主观信任作为一种认知现象, 其不确定性主要体现为模糊性, 对主观信任进行形式化研究的主要困难也在于如何对这种模糊性进行建模。

近年来, 随着对信任管理研究的深入, 出现了多种用于信任管理的模型。Wang 等人<sup>[3]</sup>提出一种基于 Bayesian 网络的信任模型, 考虑节点的不同偏好与需求, 通过更新贝塔概率函数密度的统计来计算节点信任度, 但在对节点评价的信息模糊性评判方面有欠缺。唐文等人<sup>[14]</sup>考察了信任的模糊性, 运用模糊集合论, 给出一个信任的定义评价机制, 并形式化运算规则, 但是评估机制过于复杂, 且各因素权重分配伴有随机性。

主观信任是人类的一种认知现象, 其本质是基于信念的, 而节点间信任的模糊程度也是衡量节点间信任的重要指标, 上述模型并未充分考虑这点。本文引用 Shapley 熵的概念, 提出了一种新的模糊理论的信任管理模型。当两节点交易时, 首先根据历史记录使用模糊综合评判法计算节点间直接信任。在计算推荐信任时, 考虑节点间评价取向的相似程度、节点对于评价反馈的不确定度、中间参考节点本身的可信程度三个方面。在考虑反馈评价不确定度时, 由于这种对推荐节点的不确定本身具有模糊性, 因此在计算时引用了 Shapley 熵的概念。最终综合上述三个度来计算节点推荐信任值, 并得出了综合信任值。

### 1 模型构造

在已有的信任管理模型中, 信任值一般是由直接信任值和推荐信任值组成。在本模型中对于直接信任值, 交易请求节点通过直接交易历史计算得到; 而对于推荐信任值, 则通过节点间评价取向相似度、节点反馈评价的不确定度、推荐节点的可信程度三个方面来计算, 最终得出综合信任值。

#### 1.1 相关概念

传统信息熵又称为 Shannon 熵, 是测量信息不确定性的量。信息量越大, 不确定性越小, 熵越小; 反之, 信息量越小, 不确定性越大, 熵越大。它是基于概率测度的不确定性函数, 不适合描述信息模糊的不确定性。Yager 注意到 Shapley 值的特点, 定义了基于离散模糊测度的熵, 称之为 Shapley 熵。

收稿日期: 2012-04-24; 修回日期: 2012-06-01      基金项目: 国家自然科学基金资助项目(60973146)

作者简介: 刘宇轩(1987-), 男, 硕士研究生, 主要研究方向为数学与信息安全(lawyuxuan@163.com); 郭玉翠, 女, 教授, 硕士, 主要研究方向为数学与信息安全。

**定义 1** Shapley 值或 Shapley 指标<sup>[3]</sup>。设  $U = \{u_1, u_2, \dots, u_n\}$  为有限集合,  $P(U)$  是  $U$  的幂集, 令  $\mu$  为  $P(U)$  上的模糊测度,  $\forall u_i \in U$ , 其 Shapley 值定义为

$$\Phi(u_i) = \sum_{k=0}^{n-1} \gamma_k \sum_{V \subseteq U, u_i \notin V, |V|=k} [\mu(V \cup u_i) - \mu(V)]$$

其中:  $\gamma_k = \frac{(n-k-1)! k!}{n!}$ ;  $\Phi(u_i)$  表示不确定变量  $u_i$  在整个集合  $U$  中的贡献程度, 满足  $\Phi(u_i) \in [0, 1]$ ,  $\sum_{i=1}^n \Phi(u_i) = 1$ 。

**定义 2** Shapley 熵<sup>[3]</sup>。设  $U = \{u_1, u_2, \dots, u_n\}$  为有限集合,  $P(U)$  是  $U$  的幂集, 令  $\mu$  为  $P(U)$  上的模糊测度, Shapley 熵  $H$  是关于  $\mu$  的不确定性测度:

$$H(\mu) = - \sum_{i=1}^n \Phi(u_i) \ln \Phi(u_i)$$

其中:  $\Phi(u_i)$  为上面定义的基于测度  $\mu$  的 Shapley 值。

目前, 在有关的模糊数学文献中, 多使用 Zadah 算子  $\wedge$  和  $\vee$  作为模糊算子。为了提高计算精度, 本文使用 Einstein 算子进行补充。

**定义 3** Einstein 算子  $\varepsilon, \varepsilon^+$ <sup>[13]</sup>。对于标量  $a, b$ , 有  $a \varepsilon b = \frac{a \cdot b}{1 - (1-a) \cdot (1-b)}$ ,  $a \varepsilon^+ b = \frac{a + b}{1 + a \cdot b}$ 。

对于向量  $A = (a_1, a_2, \dots, a_n)$ ,  $B = (b_1, b_2, \dots, b_n)$ , 定义

$$A \oplus B = (a_1 \varepsilon^+ b_1, \dots, a_n \varepsilon^+ b_n)$$

$$A \otimes B = (a_1 \varepsilon b_1, \dots, a_n \varepsilon b_n)$$

以及  $\bigoplus_{i=1}^n P_i = P_1 \oplus P_2 \oplus \dots \oplus P_n$ 。

## 1.2 信任值计算

在计算直接信任值时, 直接考察节点间交易历史, 通过模糊综合评价法得到结果。在计算推荐信任值时, 由各个推荐节点对目标节点的直接信任值赋权运算得出, 其中每个推荐节点所占权重不同, 在计算此推荐权重时考虑如下三个因素:

a) 评价取向相似度。交易发起节点与推荐节点之前的行为模式相似程度应影响该推荐节点的权重。现实人际交往中推荐人与顾客的行为爱好兴趣越相似, 则该推荐人对于顾客的参考价值越大。因此引入节点评价相似度, 量化两节点对于各因素的侧重相似程度。

b) 反馈信息不确定度。推荐节点对于目标节点的信任, 本身也有模糊性, 这种模糊的程度对于计算推荐权重也是重要信息。两个参考节点, 若一个对信息反馈非常确定, 而另一个模棱两可, 则前者的参考价值更大。计算这种模糊信息量, 需要引用新的数学工具 Shapley 熵。

c) 推荐节点的可信度。计算推荐信任值时, 推荐节点作为“意见参考”节点, 其本身的可信度需要考虑。对于恶意节点, 由于目标清晰, 其反馈信息不确定性很小, 必须引入可信度在计算中抑制恶意节点的参考作用。因此将推荐节点的可信度作为节点与推荐节点间的直接信任值。

得到直接信任值和推荐信任值后, 通过赋权运算得到最终的综合信任值, 作为与目标节点交易成功率的度量。

### 1.2.1 直接信任值

对于有过直接交易历史的节点  $a$  和  $b$ , 其直接信任值通过直接交易的历史经验来获得。交易完成后, 节点  $a$  对节点  $b$  的服务进行模糊评价,  $m$  种因素集合记为  $U = \{u_1, u_2, \dots, u_m\}$ ,  $n$  个评语集合记为  $V = \{v_1, v_2, \dots, v_n\}$ 。它们的元素个数和名称需要根据实际的应用情况确定, 并在此作出如下规定,  $v_1, v_2, \dots, v_n$  按照由最好评语到最差评语排列。如评语集可以取为

$$V = \{M, N, O, P\}$$

其中:  $M$  表示优,  $N$  表示良,  $O$  表示中,  $P$  表示差。

节点  $a$  与  $b$  每次交易完成后, 会对每一个因素  $u_i$  给出评语, 并更新与节点  $a$  关联的因素评判矩阵

$$R_{ab} = (r_{ij})_{m \times n} \quad (1)$$

其中:  $r_{ij}$  表示评价节点针对因素  $u_i$  给出的评价对评语  $v_j$  的频率。

每个节点对于  $U$  中的各因素侧重各有不同, 故对各因素的评价权重  $\omega^a = (\omega_1^a, \omega_2^a, \dots, \omega_m^a)$  由评价节点  $a$  给出, 满足  $\sum_{k=1}^m \omega_k^a = 1$ 。节点  $a$  对节点  $b$  的直接信任向量定义为

$$DT_{ab} = \omega^a \circ R_{ab}$$

其中:  $\circ$  运算由 Zadah 算子  $\vee$  和  $\wedge$  给出,  $DT_{ab}$  的第  $k$  个分量为

$$DT_{ab}[k] = \bigvee_{l=1}^m (\omega_l^a \wedge r_{kl})$$

得到直接信任向量后, 就可以提取出直接信任值  $dt_{ab}$ 。为防止信息流失, 运用 Einstein 算子计算, 引入衰减因子  $\lambda$ , 有

$$dt_{ab} = DT_{ab}[1] \varepsilon^+ (\lambda \varepsilon DT_{ab}[2]) \quad (2)$$

### 1.2.2 推荐信任值

节点  $a$  对节点  $b$  的信任, 一方面包含根据历史交易经验的直接信任程度; 另一方面则是其他与节点  $b$  有过直接交互的节点对  $b$  的推荐。这部分重点是计算这些推荐节点的信息参考价值, 即推荐权重。推荐权重的计算综合以下三点:

a) 评价取向相似度。对于节点  $a$ , 各项评判因素权重向量  $\omega^a = \{\omega_1^a, \omega_2^a, \dots, \omega_m^a\}$  是由  $a$  给出的, 节点  $a, c$  之间的评价取向相似度为

$$\text{sim}(a, c) = \frac{\sum_{l=1}^m \omega_l^a \omega_l^c}{\sqrt{\sum_{l=1}^m (\omega_l^a)^2} \cdot \sqrt{\sum_{l=1}^m (\omega_l^c)^2}}$$

b) 反馈信息不确定度。设推荐节点  $c$  对  $b$  的直接信任向量为

$$DT_{cb} = (t_1, t_2, \dots, t_n)$$

向量各元素中最大值记为  $t_m$ , 对向量进行处理得

$$DT_{cb}' = \left( \frac{t_1}{t_m}, \frac{t_2}{t_m}, \dots, \frac{t_n}{t_m} \right)$$

记为

$$DT_{cb}' = (t_1', t_2', \dots, t_n')$$

将  $DT_{cb}'$  的每一个分量  $t_k'$  视为  $c$  对  $b$  的评价中评语  $v_k$  的度量, 记为  $\text{tru}_c(v_k) = t_k'$ , 规定对于任意  $v \in P(V)$ ,  $V = \{v_1, v_2, \dots, v_n\}$ , 有  $\text{tru}_c(v) = \bigvee_{v_k \in v} \text{tru}_c(v_k)$ , 其中  $P(V)$  为  $V$  的幂集。

函数  $\text{tru}_c: P(V) \rightarrow [0, 1]$  满足模糊测度的两个条件:

a)  $\text{tru}_c(\emptyset) = 0, \text{tru}_c(V) = 1$ ;

b) 若  $A \subseteq B \subseteq V$ , 则  $\text{tru}_c(A) \leq \text{tru}_c(B)$ 。

故  $\text{tru}_c$  为  $V$  上的模糊测度, 现求其 Shapley 熵:

$$H(\text{tru}_c) = - \sum_{k=1}^n \Phi(v_k) \ln \Phi(v_k)$$

其中:  $\Phi(v_k)$  为定义 1 中  $v_k$  的关于  $\text{tru}_c$  的 Shapley 指标。

计算所得熵  $H(\text{tru}_c)$  在  $[0, 1]$  中, 若熵越小, 则信息量越大, 不确定性越小, 评价价值越大; 反之, 若熵越大, 则信息量越小, 不确定性越大, 评价价值越小。

c) 推荐节点的可信度。其可信度直接取节点  $a$  与推荐节点  $c$  间的直接信任值。

综合 a) ~ c), 推荐节点  $c$  的参考权重定义为

$$\mu_c^a = \frac{\text{sim}(a, c) \cdot [1 - H(\text{tru}_c)] \cdot dt_{ac}}{\sum_{c \in P} \text{sim}(a, c) \cdot [1 - H(\text{tru}_c)] \cdot dt_{ac}}$$

其中:  $P$  为所有推荐节点集合。

节点  $b$  的推荐信任向量定义为

$$RT_{ab} = \bigoplus_{c \in P} (\mu_c^a \otimes DT_{cb})$$

与式(1)类似,可得到节点  $b$  的推荐信任值计算公式为

$$rt_{ab} = RT_{ab}[1] \varepsilon^+(\lambda_2 \varepsilon RT_{ab}[2]) \quad (3)$$

### 1.2.3 综合信任值

根据式(2)(3),得出综合信任值  $t_{ab}$  为

$$t_{ab} = \alpha \cdot dt_{ab} + \beta \cdot rt_{ab} \quad (4)$$

其中: $\alpha, \beta$  分别表示直接信任值和推荐信任值对应的权值,且满足  $\alpha + \beta = 1$ 。计算得到最终的综合信任值后,节点  $a$  选取信任值最高节点进行交易。

## 2 仿真与分析

本文的仿真实验在 Windows 7 环境下用 MATLAB 7.0 实现。设定仿真网络规模为 1 000 个节点,每个节点与其他节点交易 10~50 次。在恶意节点占不同规模比例的情况下,做了本文信任模型和传统信任模型 EigenRep 的对比实验。

如图 1 所示,通过与同一实验环境下的 EigenTrust 模型进行比较,可知随着恶意节点的增多,节点间信任值倾向稳定。由于本文模型考虑了推荐信息的模糊程度,相比 EigenRep 在最终信任值计算时综合了更多信息,因此其随着恶意节点增加仍能保持较高的成功率。

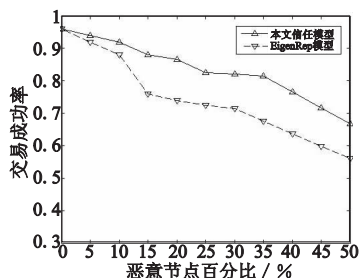


图1 交易成功率与恶意节点比率的关系

## 3 结束语

在开放网络中,信任具有的主观性及不确定性,使得精确的数学模型难以描述。本文结合模糊数学及 Shapley 熵的相关理论,提出一个基于 Shapley 熵的主观模型。其中直接信任值的计算都是根据节点的交易历史得到的;推荐信任值通过综合节点间评价取向相似度、节点反馈评价的不确定度、推荐节点的可信程度三个方面来计算,不仅考虑了推荐节点的参考信

息,同时运用 Shapley 熵准确度量了推荐信息的确定程度。最后综合直接和推荐信任值得到节点间综合信任值。交易节点的选择根据综合信任值的大小进行,综合信任值越高的节点,交易成功的可能性越大,因此节点会倾向于和信任值高的节点进行交易,从而使本模型能利用信任值来规范节点的网络行为。实验仿真表明,该模型能有效地计算节点间信任,并且对恶意行为能起到有效的抑制作用。

### 参考文献:

- [1] 苗光胜,冯登国. P2P 信任模型中基于行为相似度的共谋团体识别模型[J]. 通信学报,2009,30(8):9-20.
- [2] 林庆国,刘宴兵. 一种基于信任的动态访问控制策略[J]. 重庆邮电大学学报,2010,22(4):478-482.
- [3] 李季,朱小勇. 基于模糊理论的 Web 服务信任评估模型[J]. 计算机工程,2010,36(15):25-30.
- [4] 吕镇邦,周波. 基于 Shapley 熵和 Choquet 积分的层次化风险评估[J]. 北京邮电大学学报,2009,32(6):83-87.
- [5] 李致远,王汝传. P2P 电子商务环境下的动态安全信任管理模型[J]. 通信学报,2011,32(3):50-59.
- [6] 杨墨,王丽娜. 基于信任容错的 Web 服务可靠性增强方法研究[J]. 通信学报,2010,31(9):131-138.
- [7] 陈深龙,张玉清. 增强 Ad hoc 网络可生存性的健壮多维信任模型[J]. 通信学报,2010,31(5):1-9.
- [8] 窦文阳,王小明,张立臣. 普适环境下的动态模糊访问控制模型研究[J]. 计算机科学,2010,37(9):63-68.
- [9] 朗波. 面向分布式系统访问控制的信任度量模型[J]. 通信学报,2010,31(12):45-54.
- [10] BLAZE M, FEIGENBAUM J, LACY J. Decentralized trust management[C]//Proc of the 17th Symposium on Security and Privacy. 1996.
- [11] BETH T, BORCHERDING M, KLEIN B. Valuation of trust in open networks[C]//Proc of European Symposium on Research in Computer Security. 1994:3-18.
- [12] JØSANG A. Prospectives for modeling trust in information security[C]//Proc of Australian Conference on Information Security and Privacy. [S. l.]: Springer-Verlag, 1997.
- [13] WANG Yao, VASSILEVA J. Bayesian network-based trust model[C]//Proc of IEEE/WIC International Conference on Web Intelligence. Washington DC: IEEE Computer Society, 2003.
- [14] 唐文,陈钟. 基于模糊集合理论的主观信任管理模型研究[J]. 计算机研究与发展, 2005,42(10):1654-1659.
- [15] 窦文,王怀民,贾焰. 构造基于推荐的 Peer-to-Peer 环境下的 Trust 模型[J]. 软件学报,2004,15(4):571-583.

(上接第 4655 页)

- [3] HU Hong-gang, FENG Deng-guo. On the 2-Adic complexity and the k-Error 2-Adic complexity of periodic binary sequences[J]. IEEE Trans on Information Theory, 2008,54(2):874-883.
- [4] 王菊香,朱士信.  $F_p$  上周期序列  $S^\infty$  与  $\bar{S}^\infty$  的线性复杂度分析[J]. 计算机应用研究,2009,26(2):742-743.
- [5] NIEDERREITER H. Linear complexity and related complexity measures for sequences[C]//Lecture Notes in Computer Science, Vol 2904. Berlin: Springer-Verlag, 2003:1-17.
- [6] 王菊香,朱士信.  $F_2$  上周期多序列及其广义对偶多序列的复杂性分析[J]. 计算机应用研究,2010,27(10):3880-3882.
- [7] 王菊香,朱士信.  $P$  元周期多序列及其广义对偶多序列的复杂性分析[J]. 计算机应用研究,2011,28(10):3831-3833.
- [8] ARMAND M A. Multisequence shift register synthesis over commuta-

tive rings with identity with applications to decoding cyclic codes over integer rings[J]. IEEE Trans on Information Theory, 2004, 50(1):220-228.

- [9] CHEN Po. Multisequence linear shift register synthesis and its application to BCH decoding[J]. IEEE Trans on Communication, 1976,24(4):438-440.
- [10] FENG G L, TSENG K K. A generalized Euclidean algorithm for multisequence shift-register synthesis[J]. IEEE Trans on Information Theory, 1989,35(3):584-594.
- [11] 魏仕民,陈恺,肖国镇. 周期序列的极小多项式[J]. 西安电子科技大学学报,2000,27(6):749-751.
- [12] 王尚平,高虎明,王育民. GF(2) 上伪随机序列  $S^\infty$  与  $\bar{S}^\infty$  复杂性分析[J]. 西安电子科技大学学报:自然科学版,2002,29(1):67-70.